B39

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/12, C07K 14/705, G01N 33/68 | A2

(11) International Publication Number: WO 00/08149

(43) International Publication Date: 17 February 2000 (17.02.00)

(21) International Application Number: PCT/IB99/01445

(22) International Filing Date: 5 August 1999 (05.08.99)

(30) Priority Data:
60/095,408 5 August 1998 (05.08.98) US

(71) Applicant: CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE - CNRS [FR/FR]; 3, rue Michel Ange, F-75794 Paris Cedex 16 (FR).

(72) Inventors: WALDMANN, Rainer; Chemin de Chense, F-83600 Les Adrets de l'Esterel (FR). BASSILANA, Frédéric; 15, rue Deldille, F-06000 Nice (FR). LAZDUNSKI, Michel; 21, avenue Colombo, F-06000 Nice (FR). DE WEILLE, Jan, R.; Centre National de la Recherche Scientifique - CNRS, 3, rue Michel Ange, F-75794 Paris Cedex 16 (FR).

(74) Agents: BREESE, Pierre et al.; Breese-Majerowicz, 3, avenue de l'Opéra, F-75001 Paris (FR).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: HUMAN NEURONAL ACID-SENSITIVE CATIONIC CHANNELS, ITS CLONING AND APPLICATIONS

(57) Abstract

Non-inactivating or slowly inactivating proton-gated cation channels are thought to play an important role in the perception of pain that accompanies tissue acidosis. We have identified a human proton-gated cation channel subunit that has biphasic desensitisationkinetics with both a rapidly inactivating Na^+-selective and a sustained component. The protein shares 84 % sequence identity with the proton-gated cation channel rASIC3 (rDRASIC) from rat sensory neurones. The biphasic desnsitisation kinetics and the sequence homology suggest that this clone (hASIC3) is the human orthologue of rASIC3 (rDRASIC). While rASIC3 (rDRASIC) requires very acidic pH (< pH 4.5) for activation of the sustained current, the non-inactivating hASIC3 current starts to be activated when the pH decreases to below pH 6. hASIC3 is an acid sensor and might play an important role in the detection of lasting pH changes in human. We localized the hASIC3 gene to the human chromosome 7q35, 6.4 cRad telomeric from the microsatellite AFMA082XC9.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

HUMAN NEURONAL ACID-SENSITIVE CATIONIC CHANNELS, ITS CLONING AND APPLICATIONS

BACKGROUND OF THE INVENTION

Cross-Reference To Related Patent Applications

This patent application benefits of and claims the benefit of provisional patent application serial no. 60/095,408, *Identification, Functional Expression and Chromosomal Localisation Of Sustained Human-Proton-Gated Cation Channel*, filed on August 5, 1998, and application serial no. 09/129,758, *Mammal Neuronal Acid Sensing Cationic Channel, Cloning and Application Thereof* filed on August 5, 1998. The said U.S. applications being incorporated herein by reference in their entirety.

FIELD OF THE INVENTION

The present invention relates to new families of mammalian, notably human and rat, acidity-sensitive ionic channels. More particularly, the invention relates to the identification and molecular characterization in humans and rats of a new family of proton-activated cationic channels, collectively referred to below as ASIC polypeptides, for Acid Sensing Ionic Channel.

The ASIC channels constitute the first members of a group of cationic channels belonging to the family of amiloride-sensitive degenerine sodium channels [6, 11-14], which are activated temporarily by extracellular acidification.

Sensitivity to acid is associated with both nociception [1] and the transduction of taste [2]. The stimulation of sensory neurons by acids is of great importance because acidity accompanies numerous painful inflammatory and ischemic situations. The pain caused by acids is thought to be mediated by the cationic channels present at the level of the sensory neurons which are activated by protons [3-5]. The biophysical and pharmacological properties of the ASIC channels of the invention are similar to those of the proton-activated cationic channels described in the sensory neurons [3, 15, 16]. However, as will be seen in the description below, to date there has been no report of ligand-activated ionic channels simpler than the ASIC channels.

Summary of the Invention

The invention also relates to hybrid cationic channels constituted by the combination of a first protein comprising a proton-activated ionic channel according to the invention with a second proton-activated ionic channel.

The present invention has as its object a nucleic acid molecule coding for a protein constitutung a neuronal neuronal cationic channels that is sensitive to amiloride and activated by protons.

The invention also relates to a vector comprising at least one of the preceding nucleic acid molecules, advantageously combined with suitable control sequences, as well as a procedure for production or expression in a cell host of a protein constituting an ionic channel according to the Invention.

The invention also relates to the transformed cells expressing ASIC cation channels and/or their derivatives obtained according to the preceding methods.

The present invention also relates to application of the ASIC channel for studying pathological modifications that may lead to neuronal degenerations. The invention this also relates to the pharmaceutical preparations comprising as an active ingredient, at least one of these proteins of the invention.

Other characteristics and advantages of the invention will be seen in the description below related to research activities that led to the demonstration and the characterization of the ASIC channel.

This invention can be further understood with reference to the Figures, discussed next and in the Examples.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 represents the alignment of the sequences of the rat ASIC proteins (at top) and human ASIC proteins (at bottom) of sequences SEQ ID NO: 1 and SEQ ID NO: 2.

Figure 2 represents a comparison of the protein sequence of the rASIC1A channel with the sequence of other ionic channels:

Figure 3 represents the phylogenetic tree of the proteins of the subunits αNaCh, βNaCh, γNaCh, δNaCh of the amiloride-sensitive sodium channel and of the degenerines MEC-4, MEC-10 and DEG-1 of *C. elegans*.

Figure 4 represents the topology proposed for this latter family of ionic channels [30].

Figure 5 shows the biophysical properties of the proton-activated rASIC1A channel.

Figure 6 shows the effect of Ca^{2+} and of amiloride on the rASIC1A current.

Figure 7 shows the tissue distribution of ASIC1A channel mRNA.

Figure 8 shows the *in situ* hybridization.

Figure 9 shows the alignment of the deduced protein sequences of hASIC3 and rASIC3.

Figure 10 shows the pH dependence and pharmacology of hASIC3.

Figure 11 shows the selectivity and single channel properties of hASIC3.

Figure 12 shows the human chromosomal localization of the hASIC3 gene.

IDENTIFICATION OF THE AMINO ACID AND DNA SEQUENCES

SEQ ID NO: 1 represents the sequence of 526 amino acids of the protein of the rASIC1A channel deduced from the cDNA sequence of the rat.

SEQ ID NO: 2 represents the partial sequence of 514 amino acids of the protein of the hASIC1A channel deduced from the partial sequence of human cDNA.

SEQ ID NO: 3 represents the sequence of 512 amino acids of the protein of the hASIC2A channel deduced from the sequence of human cDNA.

SEQ ID NO: 4 represents the sequence of 559 amino acids of the protein of the rASIC1B channel as well as the sequence of a DNA molecule comprising the sequence coding for that protein.

SEQ ID NO: 5 represents the sequence of 533 amino acids of the protein of the rASIC3 channel and the sequence of DNA coding for that protein.

SEQ ID NO: 6 represents the sequence of 563 amino acids of the protein of the rASIC2B channel as well as the sequence of a DNA molecule comprising the sequence coding for that protein.

SEQ ID NO: 7 represents the sequence of 533 amino acids of the protein of the hASIC3 channel as well as the sequence of a DNA molecule comprising the sequence coding for that protein.

DETAILED DESCRIPTION OF THE INVENTION

The present invention has as its object and rat protiens constituting neuronal cationic channels that are sensitive to amiloride and which are activated by protons. The invention relates to proteins constituting the ASIC family of cation channels, or functionality equivalent derivatives of these proteins.

Such derivatives are those polypeptides whose sequence includes a modification and/or a suppression and/or an addition of one or more amino acid residues, as long as this modification, suppression, and/or addition does not alter the functional and structural properties of the ASIC channel, princi-

pally its activation by protons. Indeed, three different ASIC polypeptides, ASIC1, ASIC2 and ASIC3, in both rat and human, are described herein. In addition, the transcripts encoding ASIC1 and ASIC2 are alternatively spliced, which generates additional functional derivatives of the ASIC1 and ASIC2 proteins (ASIC1A and 1B, ASIC2A and ASIC2B, respectively). Other functional derivatives of the ASIC proteins and/or other forms of the ASIC polypeptides generated by alternative splicing of the ASIC mRNA transcripts are considered to be within the scope of the present invention. Such proteins and their functional derivatives can be analyzed by an expert in the field using the techniques described in the Examples included herein, which make it possible to demonstrate the biophysical and pharmacological properties of the ASIC channels.

Further examples of functional derivatives of the ASIC channels are as follows: The human and rat ASIC1A proteins (hASIC1A and rASIC1A, SEQ ID Nos. 1 and 2, respectively) are considered to be functionally equivalent. The amino acid sequences of these two proteins are highly homologous, but they are not identical. Thus, substitutions can readily be introduced within the primary sequence of ASIC proteins without influencing their basic functional characteristics.

Another example of such a functionally equivalent derivative is the protein constituting a cationic channel previously designated MDEG [14] or BNaCl [20], designated herein as rASIC2A. The amino acid sequence of rASIC2A is represented in the annexed list of sequences under number SEQ

ID NO: 3. rASIC2A has been described as a mammalian cationic channel which is sensitive to amiloride and which is activated in *C. elegans* by mutations that result in neurodegeneration. The rASIC2A channel is a structurally similar to the ASIC1A channel, exhibiting approximately 67% homology in their amino acid sequences. Cation transport by both polypeptides is sensitive to amiloride and regulated by acid. However, the electrophysiological properties of these two channels are different because they are not activated by the same pH changes. Thus, the range of sensitivity of rASIC2A (EC_{50} = 4.05) is different from that of ASIC1A (EC_{50} = 6.2). Other functionally equivalent proteins that may exhibit different electrophysiological properties are also considered to be within the scope of the invention.

It has been shown that the rASIC2A channel is activated by the same mutations as those causing neuronal degeneration in *C. elegans*. Thus, like the hyperactive mutants of *C. elegans*, the active mutants of rASIC2A are responsible for cell death. This indicates that the acquisition of function by this neuronal ionic channel could be associated with various forms of neuronal degeneration in mammals, notably of rodents and humans. However, no normal physiological function of rASIC2A was known until the demonstration of its activation by protons in accordance with the cationic channels of the present invention.

Other examples of proteins constituting a neuronal cationic channel that are sensitive to amiloride and activated by protons according to the invention are presented below

- A channel designated ASIC1B, whose sequence of 559 amino acids is represented in the annexed list of sequences under number SEQ ID NO: 4. ASIC1B is a splicing variant of the ASIC1A channel cloned from the rat brain by degenerated PCR. The first 185 amino acids are replaced by a new sequence of 218 amino acids which is underlined in SEQ ID NO: 4.

- A channel designated rASIC2B. rASIC2B is a splicing variant of rASIC2A and is represented by SEQ ID No. 6.

- A channel designated rASIC3, whose sequence of 533 amino acids is represented in the list of sequences under number SEQ ID NO: 5. rASIC3 was cloned from sensory neurons from the rat using a partial sequence from the data banks (Expressed Sequence Tag with accession number W62694). The properties of rASIC3 are as follows:

a) It is expressed in the sensory neurons but not in the brain.

b) Its expression in *Xenopus* oocytes or in mammalian cells allows recording of a proton-activated sodium current which presents two components: a component activating and inactivating itself rapidly, and a component activating itself more slowly and not inactivating itself. The two components are selective for Na^+. A proton-activated cationic channel that does not inactivate itself was implicated in the prolonged sensation of pain caused by acidosis.

- A channel designated hASIC3, which is represented by SEQ ID No. 7. This protein is a novel human proton-gated cation channel subunit that has biphasic desensitisation kinetics, with both a rapidly inactivating Na^+-

selective and a sustained component. The protein shares 84% sequence identity with the proton-gated cation channel rASIC3 from rat sensory neurons.

The invention also relates to hybrid cationic channels, or channels constituted by the combination of a first protein comprising a proton-activated ionic channel according to the invention with a second protein comprising a proton-activated ionic channel. Advantageously, the said second protein is also a protein comprising a proton-activated ionic channel according to the invention. An example of such a combination is illustrated by the combination of the ASIC1A, ASIC2A or ASIC3 channel with the ASIC2A channel. Such hybrid channels exhibit a third range of pH sensitivity (e.g., with ASIC: EC_{50} = 4.8). Another example of such a hybrid channel is the combination of the ASIC1A, ASIC1B, ASIC2A or ASIC3 channels with the the ASIC2B channel.

ASIC2B is a channel that was cloned from the rat brain using a partial mouse sequence accessible in the data banks (Expressed Sequence Tag with accession number W50528) and whose sequence of 563 amino acids is represented in the annexed list of sequences under number SEQ ID NO: 6. ASIC2B is a splicing variant of ASIC2A. The first 185 amino acids are replaced by a new sequence of 236 amino acids which is underlined in SEQ ID NO: 6. ASIC2B is expressed in the brain and in the sensory neurons of the dorsal root ganglia.

ASIC2B expressed alone in *Xenopus* oocytes or in mammalian cells does not form a proton-activated cationic channel. However, it can combine

with ASIC2A or ASIC3 to form proton-activated heteromultimeric channels with modified properties. The activation pH of the channel formed after the co-expression of ASIC2A and ASIC2B differs from the channel formed by ASIC2A alone. After expression of ASIC2A and ASIC2B in COS cells, the current has not reached its maximum value at pH 3 whereas the current induced by ASIC2A alone is saturated at a pH between 4.5 and 4.0. In addition, the inactivation kinetics and the ionic selectivity of the channel formed after the co-expression of ASIC2A and ASIC2B are clearly different from those of ASIC2A alone. A current appears which inactivates itself slowly and is barely selective for Na^+ and K^+.

In another example, the sodium current obtained after expression of ASIC3 becomes non-selective (it does not differentiate between sodium and potassium) when ASIC2B is co-expressed with ASIC3. This new property is similar to that of the proton-activated cationic channel which is implicated in the prolonged sensation of pain caused by acidosis. It is very probable that ASIC3 and ASIC2B are part of this channel.

The amino acid sequence homologies of the proteins constituting the ASIC1A, ASIC1B channels cited according to the invention are presented in Table 1 below.

Table 1

Channel	ASIC 1B	ASIC 1A	ASIC2B	ASIC2A	ASIC3
ASIC1B	100	80	56	61	52

ASIC1A		100	59	68	53
ASIC2B			100	78	48
ASIC2A				100	51
ASIC3					100

Polyclonal or monoclonal antibodies directed against at least one protein constituting an ion channel of the invention and/or against a hybrid channel as described above can be prepared by the classic methods described in the literature. The antibodies are useful for investigating the presence of the ionic channels of the invention in various human and animal tissues, and may also be used to inhibit or activate an ASIC channel and/or its derivatives *in vivo*. Such an application may be useful for the treatment of diseases arising from defective ASIC cation transport.

The present invention also has as its object a nucleic acid molecule coding for a protein constituting a neuronal cationic channel that is sensitive to amiloride and activated by protons. More particularly, the invention relates to a nucleic acid molecule comprising at least one sequence coding for a protein constituting the ASIC1A, ASIC1B, ASIC2A , ASIC2B, or ASIC3 cation channels from human or rat.

The invention also relates to a vector comprising at least one of the preceding nucleic acid molecules, advantageously combined with suitable control sequences, as well as a procedure for production or expression in a cell host of a protein constituting an ionic channel according to the invention. The preparation of these vectors as well as the production or expression of the

channels of the invention in a competent host cell can be accomplished by established methods known to experts in the field.

For example, the expression and production of a protein constituting a cationic channel according to the invention can be accomplished by:

- transferring a nucleic acid molecule of the invention or a vector containing said molecule into a competent host cell,
- culturing said host cell host under conditions allowing expression of the ionic channels of the invention.
- isolating the proteins constituting the ionic channels of the invention.

The host cell employed in the preceding methods can be selected from among the prokaryotes or the eukaryotes and notably from among the bacteria, yeasts or cells of mammals, plants or insects.

The vector used is selected in relation to the host to which it will be transferred; any vector such as a plasmid can be used.

The invention also relates to the transformed cells expressing ASIC cation channels and/or their derivatives obtained according to the preceding methods. These cells are useful for screening to identify substances that are capable of modulating cation transport by these polypwptides and hence, the perception of acidity with regard to both nociception and taste transduction. This screening is implemented by bringing variable quantities of a substance to be tested into contact with cells expressing the ASIC channels and determining the effects of said substance on the currents of said cation channels. These screenings allow for the identification of new drugs that are useful in the

treatment or prevention of pain. They also enable the identification and investigation of agents that modulate acid taste. In addition, these methods are useful for identifying substances that block, or can inhibit neurodegeneration induced by hyperexpression of these channels. The substances which are isolated and detected by means of the methods above are also part of the invention. The ASIC channels clearly have ionic selectivity properties, notably with regard to their selective permeability by sodium, potassium and calcium, which endows them with excitotoxic properties when hyperstimulated.

A protein constituting an ASIC neuronal ionic channel can also be useful for developing drugs intended for the treatment or prevention of pathologies entailing the painful perception of acidity which intervenes in inflammatory diseases, ischemias and a certain number of tumors. The invention thus also relates to pharmaceutical compositions comprising as active ingredients, at least one protein constituting an ionic channel according to the invention.

A nucleic acid molecule coding for a protein constituting an ASIC channel or a derivative thereof, or a vector comprising this nucleic acid molecule or a cell expressing ASIC channels are also useful for the preparation of transgenic animals. These can be animals superexpressing said channels, but also "knock-out" animals, i.e., animals deficient in the expression of these channels or of the cation transport activity of the ASIC channels. These transgenic animals are prepared by methods known to the expert in the field, and enable the

development of live models for studying animal pathologies associated with ASIC channels.

The nucleic acid molecules of the invention or the cells transformed by said molecule can thus be used for genetic therapy to compensate for a deficiency in the ASIC channels at the level of one or more tissues of a patient. The invention thus relates also to a drug comprising nucleic acid molecules of the invention or cells transformed by said nucleic acid molecules for the treatment of pathology involving the ASIC channels or their derivatives.

In addition to the property of being activated by protons and the resultant applications described above relating to the perception of acidity, the ASIC channels, and particularly ASIC channels that have genetic mutations, may be involved in some neurogenerative processes. The death of certain neurons is characteristic of many types of neuronal degenerative disorders such as Alzheimer's disease, Huntington's disease, Parkinson's disease, amyotrophic lateral sclerosis and cerebellar ataxia. Studies of such neurodegenerative processes have identified only a few deficient genes that may be responsible for or associated with these diseases. It is likely that many more important genes remain to be identified. The primitive neural network of the nematode *C. elegans* constitutes a good model of neuronal development and death. The hereditary degeneration in *C. elegans* can be due to mutations of the genes deg-1, mec-4 and mec-10. These genes exhibit homology with the subunits of amiloride-sensitive sodium channels. In addition, the functional expression of the mec-4 chimeras of the epithelial sodium channel, suggest that these genes

are ionic channels whose acquisition of function is the cause of neuronal degeneration.

The present invention thus also relates to application of the ASIC channel for studying these pathological modifications that may lead to neuronal degenerations. The techniques employed for these applications, for example for drug screening, are similar to those described above for the investigation of taste-modulating agents and analgesic agents.

In addition, a protein constituting an ASIC neuronal ionic channel, an agonist or an antagonist of said protein, can also be used for the fabrication of drugs intended for the treatment or prevention of pathologies involving cerebral neuronal degeneration. The invention thus also relates to the pharmaceutical preparations comprising as an active ingredient, at least one of these proteins of the invention, possibly combined with a physiologically acceptable vehicle.

More specifically, the invention relates to a chemical or biological substance that is capable of modifying the currents of an ionic channel and/or a hybrid channel according to the invention for the preparation of a drug capable of modulating the perception of acidity with regard to nociception as well as taste transduction in a human or animal subject.

Other characteristics and advantages of the invention will be seen in the description below related to research activities that led to the demonstration and the characterization of the ASIC channel, and in which reference will be made to the annexed sequences and drawings in which:

Detailed Description of the Figures

Figure 1 represents the alignment of the sequences of the rat ASIC proteins (at top) and human ASIC proteins (at bottom) of sequences SEQ ID NO: 1 and SEQ ID NO: 2. Comparison of these sequences shows the absence of 14 amino acids at the beginning of the human coding phase compared to that of the rat.

Figure 2 represents a comparison of the protein sequence of the rASIC1A channel with the sequence of other ionic channels:

- ASIC2A (MDEG) [14], a mammalian cationic channel that is activated by the mutations responsible for neuro-degenerations with the degenerines of *C. elegans*.

- FaNaCh [10], a peptide of a sodium channel of *Helix aspersa* that is activated by FMRFamide.

- The degenerine MEC-4 [12] of *C. elegans*.

In this figure, the residues that are identical or similar to those of ASIC are printed respectively in white on a black background and in black on a gray background. The supposed transmembranal regions (MI, MII) of rASIC1A are marked by black bars.

Figure 3 represents the phylogenetic tree of the proteins of the subunits αNaCh, βNaCh, γNaCh, δNaCh of the amiloride-sensitive sodium channel and of the degenerines MEC-4, MEC-10 and DEG-1 of *C. elegans*.

Figure 4 represents the topology proposed for this latter family of ionic channels [30].

Figure 5 shows the biophysical properties of the proton-activated rASIC1A channel.

a) the macroscopic inflowing currents recorded at -70 mV after rapid pH changes from pH 7.4 to pH 6.

b) the dose-response curve of the extracellular pH. The initial pH was 7.4 and the points represent the mean values from 6 tests. The insert in this Figure shows the typical responses at -70 mV.

c) the Q-V relations of the outside-out patch with 140 mM of Na^+ (■) or of Li^+ (●) in the bath solution. Q is the charge transported during the acid pH transition. The insert in this figure shows the typical responses in a medium containing Na^+.

d) the currents activated by the H^+ protons recorded at various potentials in an outside-out patch in a medium containing Na^+.

e) the mean i-V relations measured from the outside-out patch with 140 mM of Na^+ (■), 140 mM of Li^+ (●) or 1.8 mM of Ca^{2+} (▲), as majority permeable ions in the external solutions; the inversion potentials were respectively 65 mV, 58 mV and -34 mV.

f) the proton current through the rASIC1A channel. The relations between the current peak and the voltage were measured from an outside-out patch in a solution of free Na^+, free Ca^{2+} with pipettes containing a solution of free K^+, at pH 4 (●) and at pH 3(■), with (▲) representing the results obtained under the same conditions as (■) but with KCl in the pipette. The insert in this figure shows the typical responses under (▲) conditions.

Figure 6 shows the effect of Ca^{2+} and of amiloride on the rASIC1A current.

a) the currents activated by the H^+ protons recorded at various membranal potentials from an outside-out patch with 1.8 mM of Ca^{2+} in a solution of free Na^+; the currents were inverted at -35 mV.

b) the mean Q-V relations from an outside-out patch recorded in solutions of free Na^+ containing 1.8 mM of Ca^{2+} (o, inversion potential -34 mV) or 0.1 mM of Ca^{2+} (●, inversion potential -80 mV).

c) the effect of the external Ca^{2+} on the macroscopic peak of inflowing current recorded at - 70 mV and activated by a rapid pH change from pH 7.4 to pH 6. The insert in this Figure shows the typical responses. The points represent means values ± se of 5 oocytes.

d) the effect of amiloride on the currents activated by the H^+ protons recorded at 0 mV from an outside-out patch.

e) the inhibition of the macroscopic current (induced by a pH change from pH 7.4 to pH 6) at -70 mV by amiloride and derivatives. The points represent the means values ± se of 5 oocytes.

Figure 7 shows the tissue distribution of ASIC1A channel mRNA.

a) Northern blot analysis of the mRNA expression of the hASIC1A channel in human tissues.

b) In b: RT-PCR analysis of the mRNA expression of the rASIC1A channel in the rat brain and in the dorsal root ganglion (DRG). (+), (-) represent respectively the samples with or without reverse transcriptase. The agarose gel

sections were developed in 1% ethidium bromide. The arrows indicate the discounted size (657 pb) of the PCR product.

Figure 8 shows the *in situ* hybridization.

a,b) hybridization of 6µm sections of a dorsal root ganglion from a 3-year-old rat with the E probe marked with digoxigenin. In a: a low-lighting microphotograph (enlargement 30X). In b: a high-resolution image (enlargement 80X) of "a". One can see the intense marking of the small-diameter neurons (arrows). Similar results were also obtained with probes A, C and D.

c) the distribution of the rASIC1A channel mRNA in the brain of an adult rat analyzed by *in situ* hybridization with antisense oligonucleotide C. Identical results were obtained with oligonucleotide B. The colors indicate abundance (red: high expression; blue: not detectable). The abbreviations used in the Figure are as follows: Cer = cerebellum; Hip = hippocampus; OB = olfactory bulb; Cx = cortex.

Figure 9 shows the alignment of the deduced protein sequences of hASIC3 and rASIC3. Amino acids that are identical or similar in both sequences are printed white on black or black on grey background respectively. The two putative hydrophobic transmembrane domains are labelled with boxes. Sequences were aligned with the pileup program (Genetic Computer Group, Wisconsin).

Figure 10 shows the pH dependence and pharmacology of hASIC3. Proton-induced membrane currents were recorded from hASIC3-transfected COS cells using the whole-cell suction-pipette technique.

a) pH dependence of the hASIC3 current. H^+-gated currents were induced by decreasing the extracellular pH rapidly from pH 7.3 to the pH values indicated. The pH required for half maximal activation was pH 6.2 for the transient current and pH 4.3 for the sustained current.

b) H^+ induced hASIC3 currents depend on the resting pH. The extracellular pH was decreased rapidly from the indicated resting pH to pH 4. The currents in A and B are shown as the fraction of the saturation level of the Bolzmann fit. c) inhibition of hASIC3 by the diuretics amiloride and triamterene. In the dose-response curve for amiloride ($K_{0.5}$=15.9 μM), currents are expressed as fraction of the mean current in the absence of drug. Data points (o, transient current; l sustained current) represent the average ±SEM of at least 5 experiments. Macroscopic currents were recorded from cells clamped at -60 mV using the whole cell suction-pipette technique.

Figure 11 shows the selectivity and single channel properties of hASIC3.

a) voltage dependence of the transient and sustained whole cell current. The transient current reverses at 37.6mV, the sustained current reverses at 10.1mV.

b) the voltage dependence of the unitary currents of spontaneously active channels at pH7.3 or of channels activated by a step to pH4. Slope conductance between -10 and +40mV for both conditions is 15.0 +/- 0.6pS. V_{rev} = 30.2mV. The Na^+ equilibrium potential is at 40.1mV. Examples of spontaneous channel activity at a resting pH of 7.3 (c) or activity evoked by a drop to pH4 (d). The channel-activity recorded at pH7.3 was inhibited by 100µM

amiloride (c). Single channel currents were recorded at -60mV from outside-out membrane patches excised from hASIC transfected COS cells.

Figure 12 shows the human chromosomal localization of the hASIC3 gene. The human ASIC3 gene is localized 6.4 cRad telomeric to the framework marker AFMA082XC9 on chromosome 7 (lod score > 21). The position of hASIC3 relative to several microsatellites is shown in the right part of the Figure. The relative positions of the markers and their distances (in cRad) are the output of the RHMAPPER program. The microsatellites D7S676 and D7S642 are localized on band q35 of chromosome 7(data from http://www.ncbi.nlm.nih.gov). The cytogenic localization of those two markers is indicated with dashed lines.

Cloning the ASIC Channel

The conserved sequences of the family of ASIC ionic channels were used to prepare the following PCR primer sequences:

TTYCCIGCIRTIACIITNTGYAAY, and CAIARICCIAIITGNCCNCCDAWRTC.

A bank of rat brain cDNA (Stratagene #936515) was hybridized with the PCR product of 1 kB of rat brain and the partial clones were isolated. The fifth extremity of the cDNA (202 bp) was isolated by PCR after ligation adapted to the double-strand cDNA.

Electrophysiology

0.25 ng of cRNA was injected into the *Xenopus laevis* oocytes and the recording microelectrodes for the imposed voltage and for the patch-clamp were installed two days after the injection. The bath solutions for the outside-

out patch recordings and the pipettes for the outside-out patch and total cells recordings contained: 140 mM KCl (or NMDG), 2 mM $MgCl_2$, 5 mM EGTA, 10 mM Hepes, pH 7.4 (with KOH). The pipettes for the outside-out patch recordings and the bath solutions for the outside-out patch and total cells recordings contained: 140 mM NaCl (or LiCl or NMDGCl), 2 mM $MgCl_2$, 1.8 mM $CaCl_2$, 10 mM Hepes, pH 7.4 (adjusted with HCl, NaOH, LiOH or TMAOH). The rapid pH changes from the initial pH were obtained by perfusion with a bath solution adjusted to the pH indicated in the Figures. The intracellular acidification of the oocytes was implemented by injecting 50 ml of the internal solution at pH 2 or by perfusion and withdrawal of a bath medium containing 20 mM NH_4Cl. None of the recorded currents was contaminated by the Ca^{2+} current sensitive to the Cl^- of the *Xenopus* oocyte. The data were sampled at 2 kHz and filtered at 500 Hz for the analysis (Logiciel Biopatch).

Northern blot analysis, RT-PCR and in-situ hybridization

The Northern blot kit was obtained from Clontech Co. (Palo Alto, CA) and contained circa 2 µg of poly(A+) RNA per line. The blot was hybridized with a fragment of the partial human clone (corresponding to bases 270 to 764 of the rat clone) marked with ^{32}P at 65¡C in 6xSSC. For the RT-PCR analysis, 5 µg of rat brain total RNA and 3 µg of dorsal root ganglion were reverse transcribed and 1/30 of the sample was amplified by 30 PCR cycles with the following sequence primers:

ATTGCTCTTCCCATCTCTAT, and TTCAAGGCCCATACCTAAGT.

The negative controls were treated in an identical manner with the exception of the reverse transcriptase which was not added. The antisense oligonucleotides corresponding to base 70 to 114 (A), 215 to 248 (B), 1821 to 1859 (C), 1896 to 1940 (D) and the double-strand DNA corresponding to base 1685 to 2672 were used for the *in-situ* hybridizations. The sections of adult rat brain were hybridized with oligonucleotides B or C the ends of which were marked with ^{32}P for one night at 37¡C in 50% formamide, 2xSSC, then washed at ambient temperature in 1xSSC. The signal was eliminated by 500-times excess of unmarked oligonucleotides. The dorsal root ganglion sections were hybridized with oligonucleotides A, C or D marked with digoxigenin (DIG)-dUTP and with probe E marked with DIG-dUTP by PCR. The marking of the probes, the preparation of the samples, the hybridization and the visualization of the DIG nucleic acids with alkaline phosphatase conjugated with anti-DIG antibodies were performed in accordance with the supplier's protocols (Boehringer Mannheim).

Computer analysis

The sequence alignments and the phylogenetic tree (Kimura substitution, UPGMA option) were performed with the GCG program (Genetics Computer Group, Madison, WI).

Identification of hASIC3

Comparison of the rat DRASIC protein sequence with the database of expressed sequence tags (EST) identified two partial cDNA sequences from human total fetus (Genbank accession AA449579 and AA449322).

Both sequences originate from the same clone (IMAGE ID 785700) that we obtained from the UK HGMP RESOURCE CENTRE. Sequencing both strands using an Applied Biosystems automatic sequencer showed that the clone contains the entire coding sequence.

Chromosomal localization

The human ASIC3 gene was mapped by PCR on the Genebridge 4 Radiation Hybrid DNA panel with the primers CGATTGCAGTTCAGCATCTCT (sense) and ACCATTCGGCAGCCGCACTT (antisense) at an anealing temperature of 65°C. The PCR products were analyzed on 2% agarose gels. Samples were considered positive when a strong amplification of a 159 bp fragment was detected (Code 1), ambiguous when a faint amplification of this fragment was detected (Code 2) and negative when no amplification around 160 bp was visible (Code 0). The positive control (human genomic DNA) was positive and the negative control (hamster genomic DNA) was negative. The following code sequence for the 83 radiation hybrids was obtained and entered into the RHMAPPER program on the Whithead Institute (http://www-genome.wi.mit.edu) with a Lod score cutoff of 21: 00000 00100 00001 00021 00100 12010 00000 12112 21000 00001 10120 00010 00102 11010 00010 00212 11011 00001 100.

Expression in COS cells

The vector containing the hASIC3 coding sequence was linearized with *Not*I and blunt ended with T4 DNA polymerase. After inactivation of the T4 DNA polymerase, the hASIC3 coding sequence was excised with *Eco*RI

and subsequently subcloned into the *Eco*RI/*Sal*I (blunt) digested PCI expression vector (Promega). COS cells, at a density of 20.000 cells per 35 mm diameter petri dish, were transfected with a mix of CD8 and hASIC3-PCI (1:5) using the DEAE-Dextran method. Cells were used for electro-physiological measurements one to three days after transfection. Successfully transfected cells were recognised by their ability to fix CD8-antibody-coated beads [13].

Electrophysiology

Ion currents were recorded using either the whole cell or outside-out patch-clamp technique. The pipette solution contained (in mM): KCl 120, NaCl 30, $MgCl_2$ 2, EGTA 5, HEPES 10 (pH 7.2). The bath solution contained in mM: NaCl 140, KCl 5, $MgCl_2$ 2, $CaCl_2$ 2, HEPES 10 (pH 7.3). Changes in extracellular pH were induced by opening one out of six outlets of a microperfusion system in front of the cell or patch. Test solutions having a pH of less then 6 were buffered with 10 mM MES rather than HEPES but were identical to the control solution in all other respects. Experiments were carried out at room temperature (20-24 °C).

Results

The 35 kb cDNA isolated from rat brain codes for a protein of 526 amino acids that exhibits, as shown in Figure 2, homologies with all of the cloned members of the family of amiloride-sensitive degenerine sodium channels.

As shown in Figure 5, expression of the cRNA in the *Xenopus* oocytes induced an inflowing current activated by H^+ protons. The biophysical and pharmacological properties of the rASIC1A channel are close to those described for the proton-activated cationic channels of sensory neurons [3, 15, 16]. Reduction of the extracellular pH below a pH of 6.9 activates a rapidly rising and desensitized inflowing current (Figure 5a and b). This channel is activated by extracellular protons since, as shown in Figure 5 (c and d), application of an acid on the extracellular surface of the outside-out patch activates the channel. Intracellular acidification of oocytes and acidification of the intracellular surface of the outside-out patch does not activate the rASIC1A channel nor alter the rASIC1A current induced by the extracellular protons.

The analysis of curves I-V of Figure 5 (c and e) recorded with different extracellular cations shows that Na^+ is the majority permeable ion (simple conductance channel 14.3 pS). Like the proton-sensitive ionic channel of the sensory neurons [15, 16], the ASIC channel discriminates weakly between the cations (Figure 5c, e, f). In fact, the channel is also permeable to Li^+, K^+, Ca^{2+} and H^+ with the ratios $pNa^+/pLi^+ = 1.3$ (Figure 5c, e), $pNa^+/pK^+ = 13$ (Figure 5c, e), $pNa^+/CA^{2+} = 2.5$ (Figure 5e) and $pNa^+/H^+ = 0.8$ (Figure 5f). The permeability to Ca^{2+} of ASIC could be a voltage-independent entry path of Ca^{2+} into the cell. An inflowing current of Ca^{2+} into the cell via the ASIC channels can be detected in the absence of extracellular Na^+ (Figure 6a, b). As indicated in Figure 5 (e), the unitary conductance for Ca^{2+} was 5.2 pS. In

the presence of 140 mM of extracellular Na^+, augmentation of the concentrations of external Ca^{2+} diminished the amplitude of the current activated by the protons (Figure 6c), thereby demonstrating that Ca^{2+} inhibits the permeability to Na^+. Blockage by external Ca^{2+} is characteristic of the $I(H^+)$ of the sensory neurons [17]. The inflowing current activated by H^+ in the sensory neurons is inhibited by amiloride [18] and ethylisopropylamiloride (EIPA) [19]. As shown in Figure 6 (d, e), the rASIC1A channel exhibits the same pharmacology and is blocked in a reversible manner (Kd = 10 μM) by amiloride and its derivatives benzamil and EIPA.

In addition, the rASIC1A channel protein exhibits approximately 67% sequence homology with the degenerine ionic channel referred to as MDEG [14] or BNaCl [20], herein designated rASIC2. However, the electrophysiological properties of these two clones expressed in *Xenopus* oocytes are clearly different:

- As shown in Figure 5(a), the rASIC2 channel is not activated by the same pH changes as the rASIC1A channel.

- Substitution of the glycine residue in position 430 of rASIC2 by an acid-inhibiting amino acid such as valine or phenylalanine activates the channel [14], just as the mutation of alanine in position 704 of degenerine MEC-4 causes neurodegeneration in *C. elegans* [12]. Identical mutations of rASIC1A (glycine in position 431 replaced by valine or phenylalanine) do not lead to activity and the mutants cannot be activated by protons.

Proton-activated cationic channels have been described not only in the sensory neurons but also in the neurons of the central nervous system [21]. The tissue distribution of the expression of the mRNA of the hASIC1A channel is in agreement with this observation. As shown in Figure 7a, a 4.3-kb transcript was detected in the brain by Northern blot analysis and the PT-PCR results presented in Figure 7b show that the dorsal root ganglion expresses the rASIC1A mRNA. Figure 8 (a, b) shows that rASIC1A mRNA is well expressed by the small neurons of the dorsal root ganglion, which supports the fact that ASIC is the rapidly desensitizing proton-activated cationic channel described in the nociceptive sensory neurons. Whereas the presence of proton-activated cationic channels in the dorsal root ganglion is in agreement with their function of acidity detector in nociception, their role in the brain remains to be established. The results of *in-situ* hybridization in Figure 8c show a broad and heterogeneous expression of the rASIC1A channel mRNA. The highest levels of expression were observed in the principal olfactory bulb, the cerebral cortex, the hippocampus, the habenula, the basolateral amygdaloid nucleus and the cerebellum. The synaptic activity accompanies extracellular pH changes [22, 23] and the rapid localized pH changes in or close to the synaptic cleft are noticeably more saturated and stronger than the reported macroscopic fluctuations in the pH.

The proton-activated cationic channels are the only known ionic channels that are directly activated by a change in pH and it was envisaged that the extracellular fluctuations in pH played a neuromodulator role [23].

The expression of cationic channels in the brain supports in addition the hypothesis that the pH fluctuations are not solely a neuronal activation by a product, but even more a communications pathway in the central nervous system.

In addition to the rapidly inactivated proton-activated cationic channels, the presence has been reported in the sensory neurons of proton-activated cationic channels exhibiting slower kinetics [4, 24]. The proton-activated cationic channels probably form, like other cationic channels activated by a ligand [25, 26], a family of cationic channels in which different subunits or combinations of subunits constitute channels with diverse pharmacological and biophysical properties.

The sensation of acidity is not uniquely implicated in nociception but is also associated with the transduction of taste [2]. Acid stimulations activate the proton-activated cationic channels in the taste cells [2, 27] and amiloride inhibits the perception of acid taste [2]. Also, the physiological as well as pharmacological data indicate that rASIC1A and other members of this family are implicated in the transduction of taste. It is, in fact, especially surprising that the same class of ionic channels is associated with different facets of sensory perception:

- the amiloride-sensitive sodium channels are associated with the transduction of salty taste [2].

- the degenerines of *C. elegans* are implicated in mechanotransduction and have been proposed as forming the mechanosensitive ionic channels [28, 29].

- the ASIC family of channels are implicated in nociception and the transduction of acid taste.

Comparison of the rASIC3 sequence with the database of expressed sequence tags identified a novel human member of this ion channel family. This novel clone from a total human embryo library codes for a protein of 533 amino acids that shares the closest homology (84% identity, 87% homology) with rASIC3 (Fig. 9). The cloning of a nearly identical cDNA from human testis (hTNaC1), although without functional expression, was reported recently [14].

Expression of the novel hASIC3 clone in COS cells induced a H^+-gated cation current with kinetics very similar to that of rASIC3. When the pH is decreased rapidly from pH 7.3 to pH 5, a biphasic current is observed. A rapidly inactivating component is followed by a sustained current (Fig. 10A). These very peculiar kinetics that are also found with the rASIC3 [9] channel together with the sequence homology (84% amino acid, 82% nucleic acid identity) with rASIC3 suggest that this novel clone is the human ASIC3. We therefore call it hASIC3 (human Acid Sensing Ion Channel 3).

The pH dependence of the transient hASIC3 current ($pH_{0.5}$=6.2, Fig. 10A) is almost identical to that reported for rASIC3 ($pH_{0.5}$ = 6.5) [9]. However, the pH dependencies of the sustained rASIC3 and hASIC3 currents are

clearly different. While rASIC3 requires very acidic pH values (<pH 4.5) [9] for activation of the sustained current, the sustained hASIC3 current starts to activate when the extracellular pH decreases to below pH 6 and reaches half-maximal activity at pH 4.3 (Fig. 10A). The channel activity of hASIC3 depends, just as that of the rASIC3 channel, on the resting pH (Fig. 10B). The maximal activity of the transient hASIC3 current was observed when the resting pH was above pH 8, indicating that a fraction of the transiently activating H^+-gated cation channels are inactivated at physiological pH. Half-maximal activation of the transient current was observed at pH 7.5, a slightly more alkaline pH than that reported for the rASIC3 clone (pH 6.5) [9]. When the resting pH was below pH 7, only activation of the sustained current could be observed after acidification of the bath medium (Fig. 10B). The sustained hASIC3 current can, just as the sustained rASIC3 channel, still be activated when the initial pH is quite acidic (pH5) (Fig. 10B).

All members of the ASIC family cloned so far are sensitive to the diuretic amiloride. The hASIC3 channel is no exception. The effect of amiloride on the hASIC3 current is similar to that reported for rASIC3 [9]. The transient current is inhibited by amiloride (K_D=15.9 μM; Fig. 2C) as well as by triamterene (Fig. 10C), while the sustained hASIC3 current is virtually not affected by those diuretics.

The transient hASIC3 current reverses at 37.6 mV, close to the Na^+ reversal potential, indicating a high selectivity for Na^+ *vs* K^+ (Fig. 11A). Conversely, the sustained current discriminates much less between Na^+ and

K^+ (selectivity ratio gNa^+/gK^+ = 1.62) as it reverses at 10.1 mV (Fig. 11A). The low selectivity for Na^+ *vs* K^+ of the sustained hASIC3 current clearly distinguishes the hASIC3 channel from the rASIC3 channel which is highly selective for Na^+ [9].

Proton-induced unitary currents were recorded from excised outside-out patches (Fig. 11B-D). In a narrow pH window around pH 7.3, spontaneous channel activity can be observed (Fig. 11C) that disappears upon an increase in pH to 8.0, a decrease in pH to 6.0 (not shown) or in the presence of 100μM amiloride (Fig. 11C). This basal current is mainly carried by Na^+, since it reverses at 30.2 mV (Fig. 11B). When the pH on the extracellular face of an outside-out patch is decreased from pH 7.3 to pH 4, unitary currents are induced (Fig. 11D) that reverse at the same membrane potential as the spontaneously active channel (Fig. 11B). The unitary conductance of the hASIC3 channel for Na^+ is 15 ±0.6 pS, close to that reported for rat ASIC3 (12.6 pS) [9]. While the sustained non-selective H^+-activated hASIC3 current could be easily detected in whole cell recordings, no sustained or non-selective current could be recorded on outside-out patches. One possible explanation is, that soluble factors might be necessary that are lost during excision of the patch.

The human chromosomal localization of the hASIC3 gene was determined by PCR on a human-hamster radiation hybrid DNA panel. The hASIC3 gene is localised on the human chromosome 7q35, 6.4 cRad telomeric from the microsatellite AFMA082XC9 (Lod score > 21). To our

knowledge, no hereditary diseases with symptoms that are consistent with an altered function of a H^+-gated cation channel were mapped to this region of the human genome.

The hASIC3 channel subunit forms a sustained H^+-gated cation channel that has properties similar to those reported for the rASIC3 channel. However, very important differences exist. Most importantly, the sustained hASIC3 current requires less acidic pH for activation than rASIC3 [9]. In this respect the properties of the hASIC3 channel match better the physiological and electrophysiological data from sensory neurones than those of rASIC3. Subcutaneous perfusion of human volunteers with acidic buffer causes pain. At pH 5.2, the pain was rated 20% on a scale ranging from 0 to 100% (unbearable pain) [2]. Furthermore, a subpopulation of polymodal C-fibres in rat nerve-skin preparations can be excited by acidic pH [4]. The threshold for activation lies between pH 6.9 and pH 6.1, maximal stimulation is reached at pH 5.2. The endogenous H^+-gated cation channel recorded in rat sensory neurones starts to activate below pH 6.6 [5]. The pH dependence of the sustained hASIC3 current matches closely those physiological data, while rASIC3 has a pH dependence that is shifted two pH units towards more acidic pH values [9]. One possible explanation for the differences between physiological data and the pH dependence of the sustained ASIC3 channel (especially the rASIC3) might be the participation of as yet unknown subunits in the formation of the native channel. Heteromultimeric assembly was previously demonstrated for the rASIC3 channel [9]. rASIC3 can associate

with rASIC2b resulting in an altered selectivity of the channel. While rASIC3 is completely Na^+-selective, the sustained current of the heteromultimeric rASIC3/rASIC2b channel does not discriminate between Na^+ and K^+. The H^+-gated cation channel recorded in rat sensory neurones does not discriminate between Na^+ and K^+ either [5], suggesting that both rASIC3 and rASIC2b participate in the formation of this ion channel in rat sensory neurons. In contrast with the rASIC3 channel, hASIC3 does not require coexpression of other subunits to generate a non-selective sustained current. The ion selectivity of sustained human H^+-gated cation channels is not known yet. A more detailed electrophysiological characterization of human sustained H^+-gated cation channels will be necessary to allow a comparison of the properties of the native channel with those of the hASIC3 channel.

List of Sequences

INFORMATION CONCERNING SEQ ID NO: 1

i) CHARACTERISTIC OF THE SEQUENCE:

A) LENGTH: 3562 base pairs

B) TYPE: nucleic acid

C) NUMBER OF STRANDS: double

D) CONFIGURATION: linear

ii) TYPE OF MOLECULE: DNA

vi) ORIGIN: rat

ix) CHARACTERISTIC

A) NAME/KEY: ASIC

B) LOCALIZATION: 123 .. 1700

xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO: 1:

Top of page 28 =

INFORMATION CONCERNING SEQ ID NO: 2

i) CHARACTERISTIC OF THE SEQUENCE:

A) LENGTH: 1620 base pairs

B) TYPE: nucleic acid

C) NUMBER OF STRANDS: double

D) CONFIGURATION: linear

ii) TYPE OF MOLECULE: DNA

vi) ORIGIN: human

ix) CHARACTERISTIC

A) NAME/KEY: ASIC

B) LOCALIZATION: 1 .. 1542

xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO: 2:

Top of page 31 =

INFORMATION CONCERNING SEQ ID NO: 3

i) CHARACTERISTIC OF THE SEQUENCE:

A) LENGTH: 1666 base pairs

B) TYPE: nucleic acid

C) NUMBER OF STRANDS: double

D) CONFIGURATION: linear

ii) TYPE OF MOLECULE: DNA

vi) ORIGIN: human

ix) CHARACTERISTIC

A) NAME/KEY: MDEG

B) LOCALIZATION: 127 .. 1663

xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO: 3:

Top of page 34 =

INFORMATION CONCERNING SEQ ID NO: 4

i) CHARACTERISTIC OF THE SEQUENCE:

A) LENGTH: 3647 base pairs

B) TYPE: nucleic acid

C) NUMBER OF STRANDS: double

D) CONFIGURATION: linear

ii) TYPE OF MOLECULE: DNA

vi) ORIGIN: rat

ix) CHARACTERISTIC

A) NAME/KEY: ASIC1B

B) LOCALIZATION: 109 .. 1785

xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO: 4:

Top of page 38 =

INFORMATION CONCERNING SEQ ID NO: 5

i) CHARACTERISTIC OF THE SEQUENCE:

A) LENGTH: 1602 base pairs

B) TYPE: nucleic acid

C) NUMBER OF STRANDS: double

D) CONFIGURATION: linear

ii) TYPE OF MOLECULE: DNA

vi) ORIGIN: rat

ix) CHARACTERISTIC

A) NAME/KEY: ASIC3

B) LOCALIZATION: 1 .. 1602

xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO: 5:

Top of page 41 =

INFORMATION CONCERNING SEQ ID NO: 6

i) CHARACTERISTIC OF THE SEQUENCE:

A) LENGTH: 1948 base pairs

B) TYPE: nucleic acid

C) NUMBER OF STRANDS: double

D) CONFIGURATION: linear

ii) TYPE OF MOLECULE: DNA

vi) ORIGIN: rat

ix) CHARACTERISTIC

A) NAME/KEY: ASIC2B

B) LOCALIZATION: 16 .. 1707

xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO: 6:

INFORMATION CONCERNING SEQ ID NO: 7

i) CHARACTERISTIC OF THE SEQUENCE:

A) LENGTH: 1736 base pairs

B) TYPE: nucleic acid

C) NUMBER OF STRANDS: double

D) CONFIGURATION: linear

ii) TYPE OF MOLECULE: DNA

vi) ORIGIN: human

ix) CHARACTERISTIC

A) NAME/KEY: ASIC3

B) LOCALIZATION: 18 .. 1611

xi) DESCRIPTION OF THE SEQUENCE: SEQ ID NO: 7:

INFORMATION CONCERNING SEQ ID NO: 8

i) CHARACTERISTIC OF THE SEQUENCE:

A) LENGTH: 531

B) TYPE: protein

C) NUMBER OF STRANDS: single

D) CONFIGURATION: linear

ii) TYPE OF MOLECULE: protein

vi) ORIGIN: human

ix) CHARACTERISTIC

A) NAME/KEY: hASIC3

B) LOCALIZATION: 1 - 531

REFERENCES

1. Rang, H.P., Bevan, S. & Dray, A. *Br. Med. Bull.* 47, 534-548 (1991).

2. Lindeman, B. *Physiol. Rev.* 76, 718-766 (1996).

3. Krishtal, O.A. & Pidoplichko, V.I. *Neuroscience* 6, 2599-2601 (1981).

4. Bevan, S. & Geppeti, P. *Trends Neurosci.* 17, 509-512 (1994).

5. Akaike, N., Krishtal, O.A. & Maruyama, T. *J. Neurophysiol.* 63, 805-813 (1990).

6. Canessa, C.M., Horisberger, J.D. & Rossier, B.C. *Nature* 361, 467-470 (1993).

7. Canessa, C.M., Schild, L., Buell, G., Thorens, B., Gautschi, I., Horisberger, J.D. & Rossier, B.C. *Nature* 367, 463-467 (1994).

8. Lingueglia, E., Voilley, N., Waldmann, H., Lazunski, M. & Barbry, P. *Febs Lett. 318, 95-99 (1993).*

9. Lingueglia, E., Renard, S., Waldmann, R., Voilley, N., Champigny, G., Plass, H., Lazunski, M. & Barbry, P., *J. Biol. Chem.* 269, 13736-13739 (1994)..

10. Lingueglia, E., Champigny, G., Lazdunski, M. & Barbry, P. *Nature* 378, 730-733 (1995).

11. Waldmann, R., Champigny, G., Bassilana, F., Voilley, N. & Lazdunski, M. *J. Biol. Chem.* 270, 27411-27414 (1995).

12. Driscoll, M. & Chalfie, M. *Nature* 349, 588-593 (1991).

13. Huang, M. & Chalfie, M. *Nature* 367, 467-470 (1994).

14. Waldmann, R., Champigny, G., Voilley, N., Lauritzen, I. & Lazdunski, M. *J. Biol. Chem.* 271, 10433-10434 (1996).

15. Kovalchuk Yu, N., Krishtal, O.A. & Nowycky, M.C. *Neurosci. Lett.* 115-237-242 (1990).

16. Konnerth, A., Lux, H.D. & Morad, M. *J. Physiol.* 386, 603-633 (1987).

17. Davies, N.W., Lux, H.D. & Morad, M. *J. Physiol.* 400, 159-187 (1988).

18. Korkushko, A.O. & Krishtal, O.A. *Neirofiziologiia* 16, 557-561 (1984).

19. Grantyn, R., Perouansky, M., Rodriguez-Tebar, A. & Lux, H.D. *Dev. Brain Res.* 49, 150-155 (1989).

20. Price, M.P., Snyder, P.M. & Welsh, M.J. *J. Biol. Chem.* 271, 7879-7882 (1996).

21. Akaike, N. & Ueno, *S. Prog. Neurobiol.* 43, 73-83 (1994).

22. Krishtal, O.A., Osipchuk, Y.V., Shelest, T.N. & Smirnoff, S.V. *Brain Res.* 436, 352-356 (1987).

23. Chesler, M. & Kaila, K. *Trends Neurosci.* 15, 396-402 (1992).

24. Bevan, S. & Yeats, J. *J. Physiol.* 433, 145-161 (1991).

25. Lewis, C., Neidhart, S., Holy, C., North, R.A., Buell, G. & Surprenant, A. *Nature* 377, 432-435 (1995).

26. Barnard, E.A. *Trends Pharmacol. Sci.* 17, 305-309 (1996).

27. Okada, Y., Miyamoto, T. & Sato, T. *J. Exp. Biol.* 187, 19-32 (1994).

28. Liu, J. Schrank, B. & Waterson, R. *Science* 273, 361 (1996).

29. Waldmann, R., Champigny, G. & Lazdunski, M. *J. Biol. Chem.* 270, 11735-11737 (1995).

30. Renard, S., Lingueglia, E., Voilley, N., Lazdunski, M. & Barbry, P. *J. Biol. Chem.* 269, 12981-12986 (1994).

31. Reeh, P.W. and Steen, K.H. *Prog Brain Res* 113, 143-151 (1996).

32. Steen, K.H., Steen, A.E., Kreysel, H.W. & Reeh, P.W. *Pain* 66, 163-170 (1996).

33. Steen, K.H., Issberner, U. & Reeh, P.W. *Neurosci Lett* 199, 29-32 (1995).

34. Steen, K.H., Reeh, P.W., Anton, F.& Handwerker, H.O. *J. Neurosci* 12, 86-95 (1992).

35. Waldmann, R., Champigny, G., Bassilana, F., Heurteaux, C. and Lazdunski, M. *Nature* 386, 173-177 (1997).

36. Bassilana, F., Champigny, G., Waldmann, R., de Weille, J.R., Heurteaux, C. & Lazdunski, M. *J. Biol. Chem.* 272, 28819-28822 (1997).

37. Waldmann, R., Bassilana, F., de Weille, J., Champigny, G., Heurteaux, C. & Lazdunski, M. *J. Biol. Chem.* 272, 20975-20978 (1997)

38. Lingueglia, E., de Weille, J.R., Bassilana, F., Heurteaux, C., Sakai, H., Waldmann, R. & Lazdunski, M. *J. Biol Chem.* 272, 29778-29783 (1997).

39. Waldmann, R. & Lazdunski, M. *Curr. Op. Neurobiol.* 8, 418-424 (1998).

40. Baumann, T.K., Burchiel, K.J., Ingram, S.L. & Martenson, M.E. *Pain* 65, 31-38 (1996).

41. Jurman, M.E., Boland, L.M., Liu, Y. & Yellen, G. *Biotechniques* 17, 876-881 (1994).

42. Ishibashi, K. & Marumo, F. *Biochem. Biophys Res Comm* 245, 589-593 (1998).

We claim:

1. An isolated and purified nucleic acid molecule encoding the human cation transport protein ASIC3.

2. The sequence of claim 1 which maps to chromosome 7q35.

3. The sequence of claim 1 represented by SEQ ID No. 7.

4. An isolated and purified human cation transport protein represented by SEQ ID No. 8.

5. The protein of claim 4 which comprises a proton gated cation channel.

6. The protein of claim 5 wherein the transport channel activity is dependent upon resting pH.

7. The protein of claim 5 which also exhibits biphasic desensitization kinetics.

8. The protein of claim 4 which is sensitive to amiloride.

9. A human cation transport channel comprising at least one molecule of ASIC3 and at least one molecule of a second human proton-gated cation channel.

10. The cation transport channel of claim 9 wherein the second cation transport protein is selected from the group consisting of hASIC2A, hASIC2B, hASIC1A and hASIC1B.

11. A cloning vehicle comprising the sequence of claim 1.

12. The cloning vehicle of claim 11 which also comprises a sequence encoding a second cation transport protein selected from the group consisting of hASIC2A, hASIC2B, hASIC1A and hASIC1B.

13. A transformed cell containing the cloning vehicle of claim 11.

14. The transformed cell of claim 13 which expresses hASIC3.

15. A transformed cell containing the cloning vehicle of claim 12.

16. The transformed cell of claim 15 which expresses a human cation transport channel comprising at least one molecule of ASIC3 and at least one molecule of a a second cation transport protein selected from the group consisting of hASIC2A, hASIC2B, hASIC1A and hASIC1B.

17. A method of screening for substances capable of modulating the activity of cation transport channels comprised of hASIC3, comprising contacting pre-selected amounts of the substance to be tested with cells expressing said cation transport channel, measuring the effects of the substance on the transport functions of the cation transport channel, and identifying the substance that has a positive or negative effect on potassium channel activity .

18. A substance, identified by the method of claim 20 that is capable of positively or negatively influencing the transport functions of a cation transport channel.

Met Glu Leu Lys Thr Glu Glu Glu Glu Val Gly Gly Val Gln Pro Val Ser Ile
Pro Val Ser Ile

Gln Ala Phe Ala Ser Ser Ser Thr Leu His Gly Leu Ala His Ile Phe Ser Tyr
Gln Ala Phe Ala Ser Ser Ser Thr Leu His Gly Met Ala His Ile Phe Ser Tyr

Glu Arg Leu Ser Leu Lys Arg Ala Leu Trp Ala Leu Cys Phe Leu Gly Ser Leu
Glu Arg Leu Ser Leu Lys Arg Ala Leu Trp Ala Leu Cys Phe Leu Gly Ser Leu

Ala Val Leu Leu Cys Val Cys Thr Glu Arg Val Gln Tyr Tyr Phe Cys Tyr His
Ala Val Leu Leu Cys Val Cys Thr Glu Arg Val Gln Tyr Tyr Phe His Tyr His

His Val Thr Lys Leu Asp Glu Val Ala Ala Ser Gln Leu Thr Phe Pro Ala Val
His Val Thr Lys Leu Asp Glu Val Ala Ala Ser Gln Leu Thr Phe Pro Ala Val

Thr Leu Cys Asn Leu Asn Glu Phe Arg Phe Ser Gln Val Ser Lys Asn Asp Leu
Thr Leu Cys Asn Leu Asn Glu Phe Arg Phe Ser Gln Val Ser Lys Asn Asp Leu

Tyr His Ala Gly Glu Leu Leu Ala Leu Leu Asn Asn Arg Tyr Glu Ile Pro Asp
Tyr His Ala Gly Glu Leu Leu Ala Leu Leu Asn Asn Arg Tyr Glu Ile Pro Asp

Thr Gln Met Ala Asp Glu Lys Gln Leu Glu Ile Leu Gln Asp Lys Ala Asn Phe
Thr Gln Met Ala Asp Glu Lys Gln Leu Glu Ile Leu Gln Asp Lys Ala Asn Phe

Arg Ser Phe Lys Pro Lys Pro Phe Asn Met Arg Glu Phe Tyr Asp Arg Ala Gly
Arg Ser Phe Lys Pro Lys Pro Phe Asn Met Arg Glu Phe Tyr Asp Arg Ala Gly

His Asp Ile Arg Asp Met Leu Leu Ser Cys His Phe Arg Gly Glu Ala Cys Ser
His Asp Ile Arg Asp Met Leu Leu Ser Cys His Phe Arg Gly Glu Val Cys Ser

Ala Glu Asp Phe Lys Val Val Phe Thr Arg Tyr Gly Lys Cys Tyr Thr Phe Asn
Ala Glu Asp Phe Lys Val Val Phe Thr Arg Tyr Gly Lys Cys Tyr Thr Phe Asn

Ser Gly Gln Asp Gly Arg Pro Arg Leu Lys Thr Met Lys Gly Gly Thr Gly Asn
Ser Gly Arg Asn Gly Arg Pro Arg Leu Lys Thr Met Lys Gly Gly Thr Gly Asn

Gly Leu Glu Ile Met Leu Asp Ile Gln Gln Asp Glu Tyr Leu Pro Val Trp Gly
Gly Leu Glu Ile Met Leu Asp Ile Gln Gln Asp Glu Tyr Leu Pro Val Trp Gly

Glu Thr Asp Glu Thr Ser Phe Glu Ala Gly Ile Lys Val Gln Ile His Ser Gln
Glu Thr Asp Glu Thr Ser Phe Glu Ala Gly Ile Lys Val Gln Ile His Ser Gln

Asp Glu Pro Pro Phe Ile Asp Gln Leu Gly Phe Gly Val Ala Pro Gly Phe Gln
Asp Glu Pro Pro Phe Ile Asp Gln Leu Gly Phe Gly Val Ala Pro Gly Phe Gln

Figure 1

Thr Phe Val Ser Cys Gln Glu Gln Arg Leu Ile Tyr Leu Pro Ser Pro Trp Gly
Thr Phe Val Ala Cys Gln Glu Gln Arg Leu Ile Tyr Leu Pro Pro Pro Trp Gly

Thr Cys Asn Ala Val Thr Met Asp Ser Asp Phe Phe Asp Ser Tyr Ser
Thr Cys Lys Ala Val Thr Met Asp Ser Asp Leu Asp Phe Phe Asp Ser Tyr Ser

Ile Thr Ala Cys Arg Ile Asp Cys Glu Thr Arg Tyr Leu Val Glu Asn Cys Asn
Ile Thr Ala Cys Arg Ile Asp Cys Glu Thr Arg Tyr Leu Val Glu Asn Cys Asn

Cys Arg Met Val His Met Pro Gly Asp Ala Pro Tyr Cys Thr Pro Glu Gln Tyr
Cys Arg Met Val His Met Pro Gly Asp Ala Pro Tyr Cys Thr Pro Glu Gln Tyr

Lys Glu Cys Ala Asp Pro Ala Leu Asp Phe Leu Val Glu Lys Asp Gln Glu Tyr
Lys Glu Cys Ala Asp Pro Ala Leu Asp Phe Leu Val Glu Lys Asp Gln Glu Tyr

Cys Val Cys Glu Met Pro Cys Asn Leu Thr Arg Tyr Gly Lys Glu Leu Ser Met
Cys Val Cys Glu Met Pro Cys Asn Leu Thr Arg Tyr Gly Lys Glu Leu Ser Met

Val Lys Ile Pro Ser Lys Ala Ser Ala Lys Tyr Leu Ala Lys Lys Phe Asn Lys
Val Lys Ile Pro Ser Lys Ala Ser Ala Lys Tyr Leu Ala Lys Lys Phe Asn Lys

Ser Glu Gln Tyr Ile Gly Glu Asn Ile Leu Val Leu Asp Ile Phe Phe Glu Val
Ser Glu Gln Tyr Ile Gly Glu Asn Ile Leu Val Leu Asp Ile Phe Phe Glu Val

Leu Asn Tyr Glu Thr Ile Glu Gln Lys Lys Ala Tyr Glu Ile Ala Gly Leu Leu
Leu Asn Tyr Glu Thr Ile Glu Gln Lys Lys Ala Tyr Glu Ile Ala Gly Leu Leu

Gly Asp Ile Gly Gly Gln Met Gly Leu Phe Ile Gly Ala Ser Ile Leu Thr Val
Gly Asp Ile Gly Gly Gln Met Gly Leu Phe Ile Gly Ala Ser Ile Leu Thr Val

Leu Glu Leu Phe Asp Tyr Ala Tyr Glu Val Ile Lys His Arg Leu Cys Arg Arg
Leu Glu Leu Phe Asp Tyr Ala Tyr Gly Val Ile Lys His Lys Leu Cys Arg Arg

Gly Lys Cys Gln Lys Glu Ala Lys Arg Ser Ser Ala Asp Lys Gly Val Ala Leu
Gly Lys Cys Gln Lys Glu Ala Lys Arg Ser Ser Ala Asp Lys Gly Val Ala Leu

Ser Leu Asp Asp Val Lys Arg His Asn Pro Cys Glu Ser Leu Arg Gly His Pro
Ser Leu Asp Asp Val Lys Arg His Asn Pro Cys Glu Ser Leu Arg Gly His Pro

Ala Gly Met Thr Tyr Ala Ala Asn Ile Leu Pro His His Pro Ala Arg Gly Thr
Ala Gly Met Thr Tyr Ala Ala Asn Ile Val Pro His His Pro Ala Arg Gly Thr

Phe Glu Asp Phe Thr Cys
Phe Glu Asp Phe Thr Cys

Fig 1 (continued)

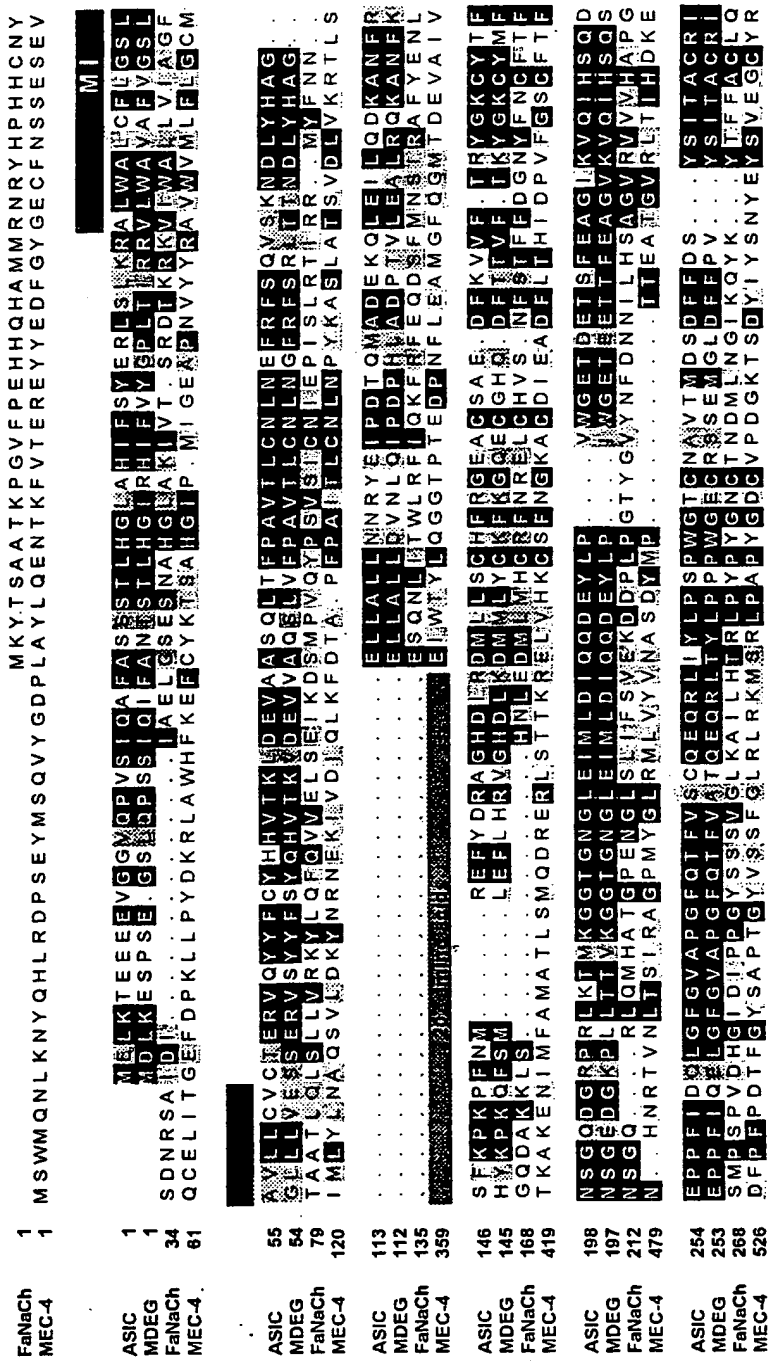


Figure 2

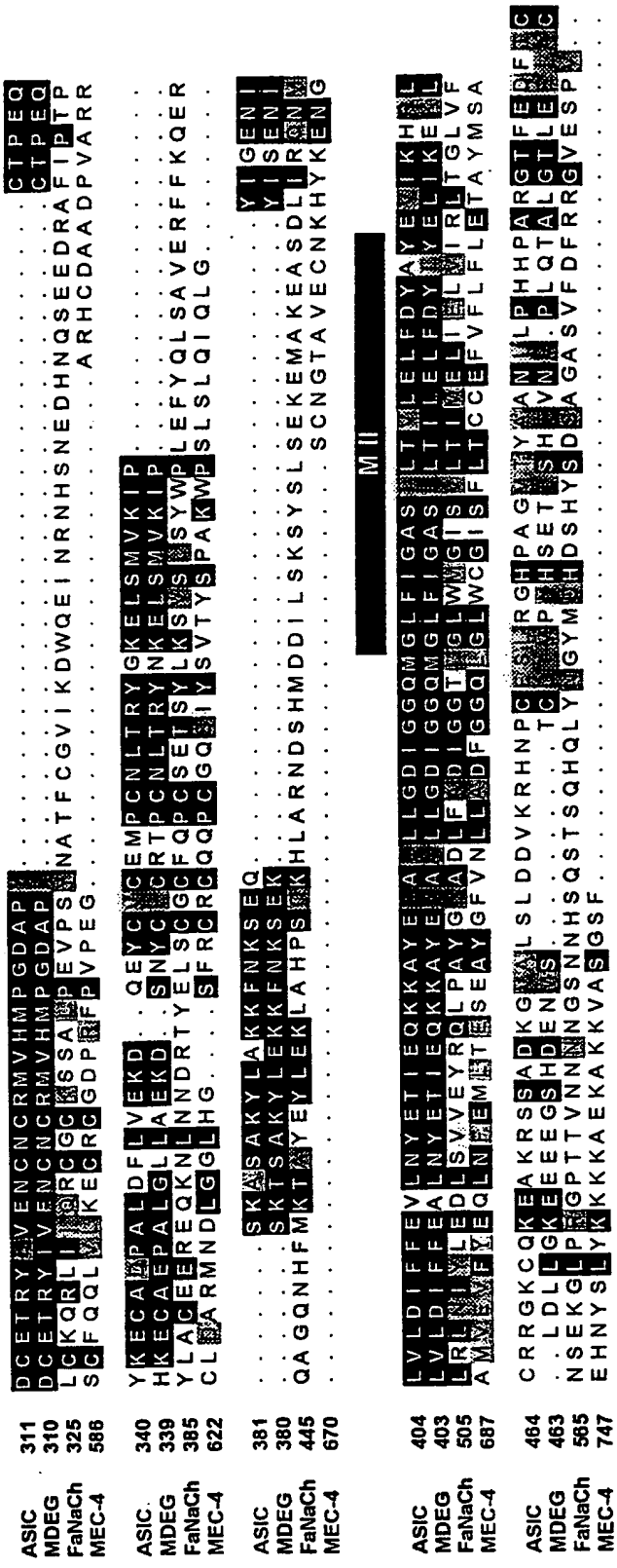


Fig.2 (continued)

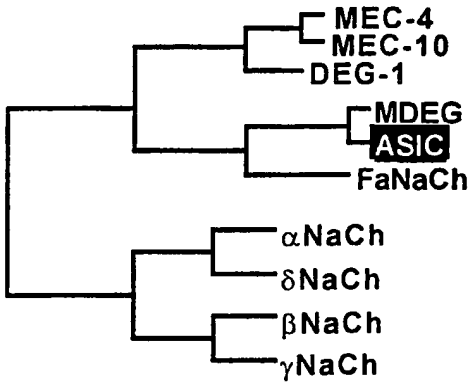


Figure 3

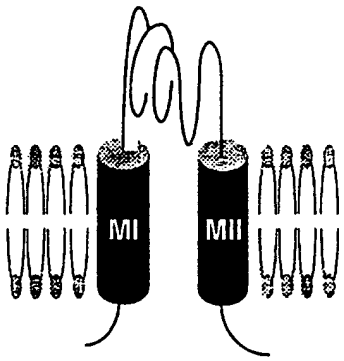


Figure 4

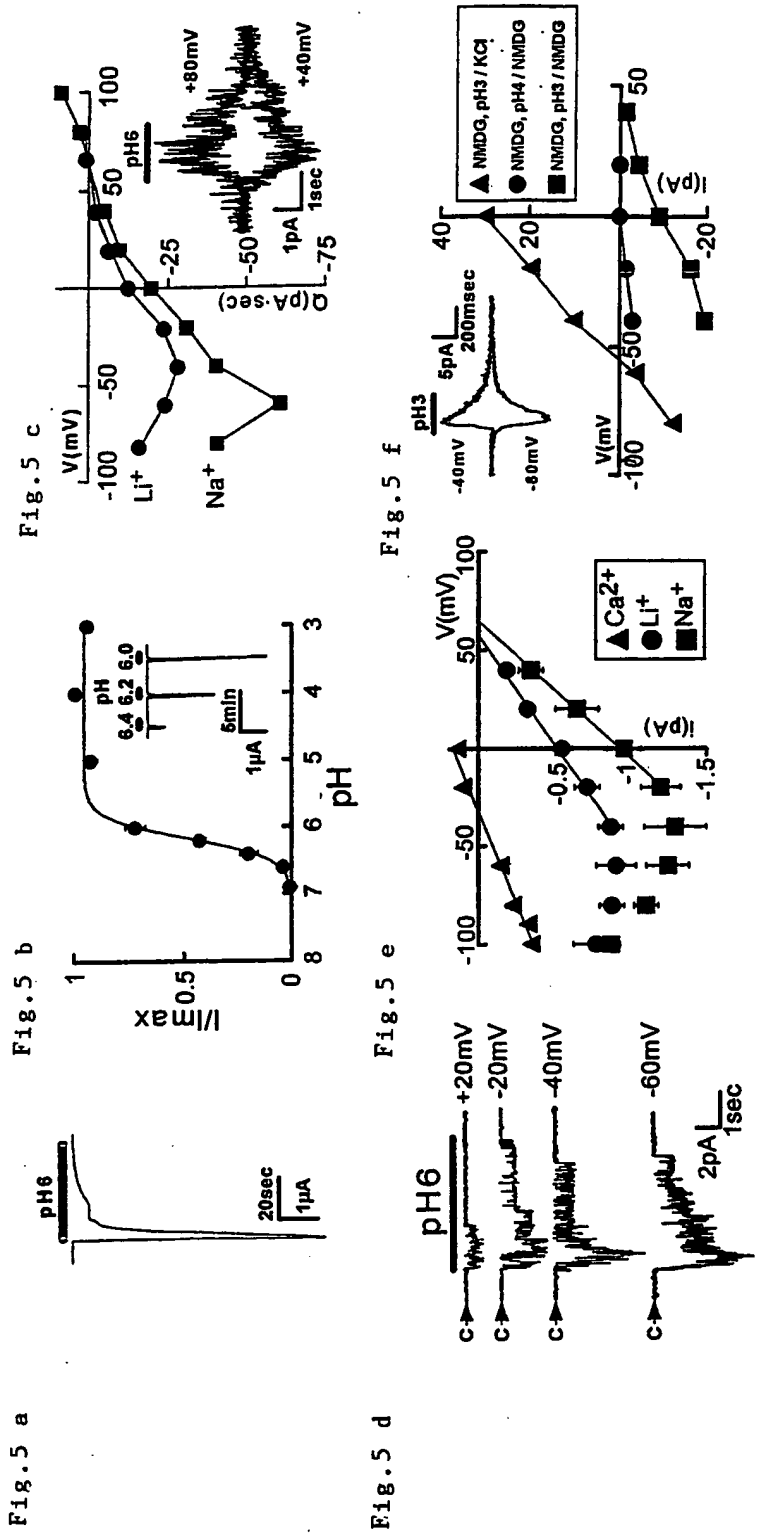


Fig.5 a
pH6
20sec
1μA
Fig.5 b
I/Imax
pH
6.4 6.2 6.0
5min
1μA
Fig.5 c
V(mV)
-100 -50 50 100
Li+
Na+
pH6
+80mV
+40mV
-25 -50 -75
Q (pA·sec)
1pA
1sec
Fig.5 d
pH6
+20mV
-20mV
-40mV
-60mV
2pA
1sec
Fig.5 e
V(mV)
-100 -50 50 100
Ca2+
Li+
Na+
i(pA)
-0.5 -1 -1.5
Fig.5 f
NMDG, pH3 / KCl
NMDG, pH4 / NMDG
NMDG, pH3 / NMDG
pH3
5pA
200msec
-40mV
-80mV
V(mV)
-100 -50 50
i(pA)
40 20 -20

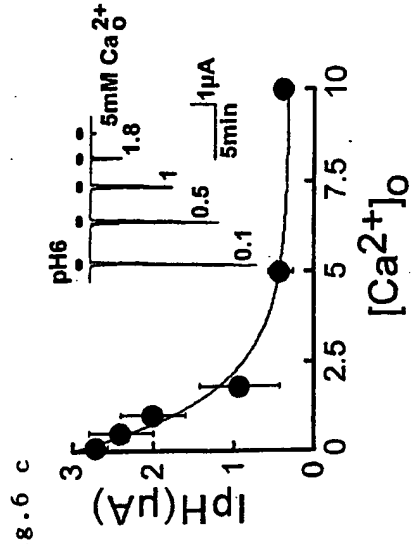


Fig.6 c

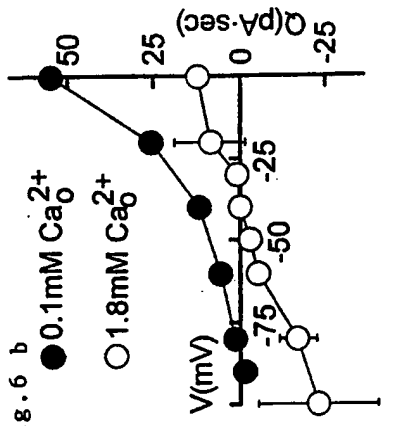


Fig.6 b

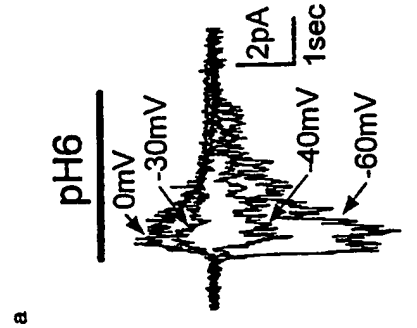


Fig.6 a

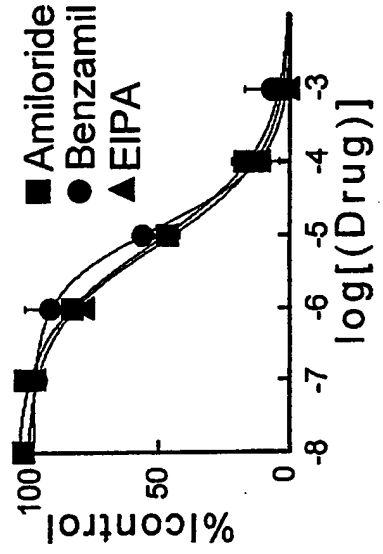


Fig.6 e

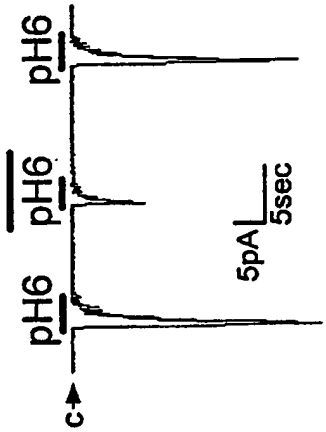


Fig.6 d

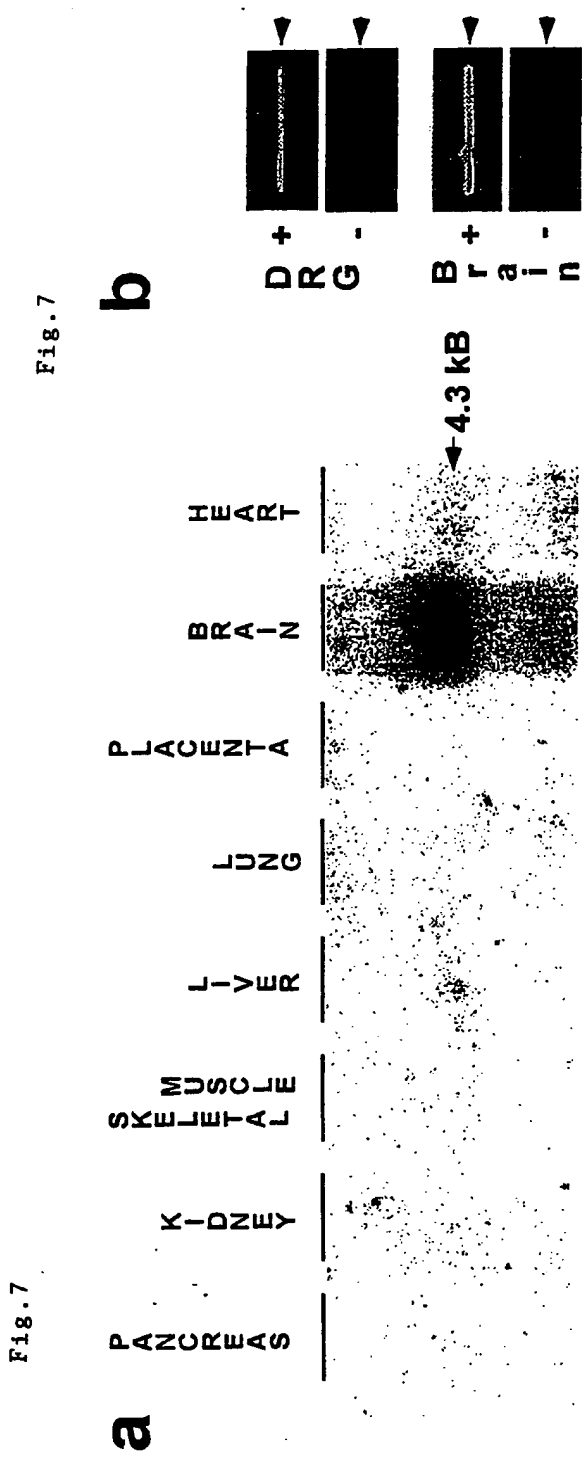

b
D+
R-
G+
Brain-
4.3 kB
a
PANCREAS
KIDNEY
SKELETAL MUSCLE
LIVER
LUNG
PLACENTA
BRAIN
HEART

Fig.7

Fig.7

Fig.8 a

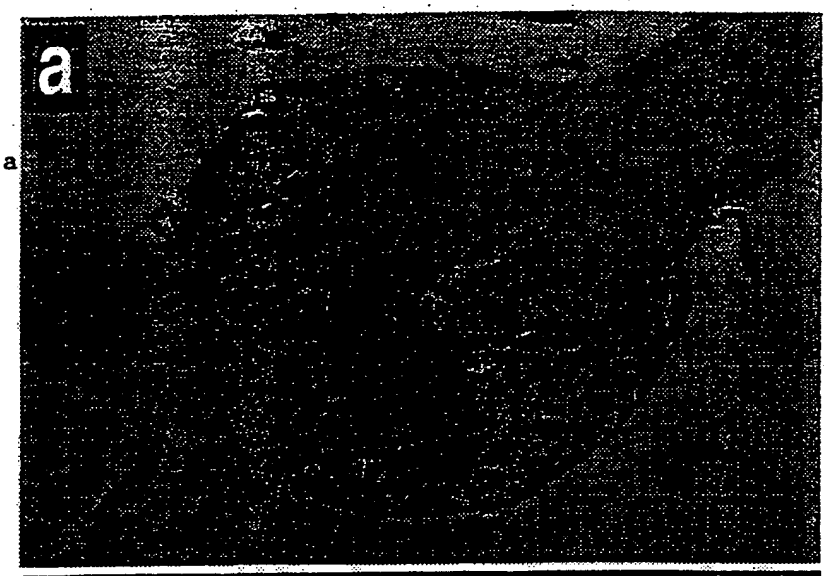


Fig.8 b

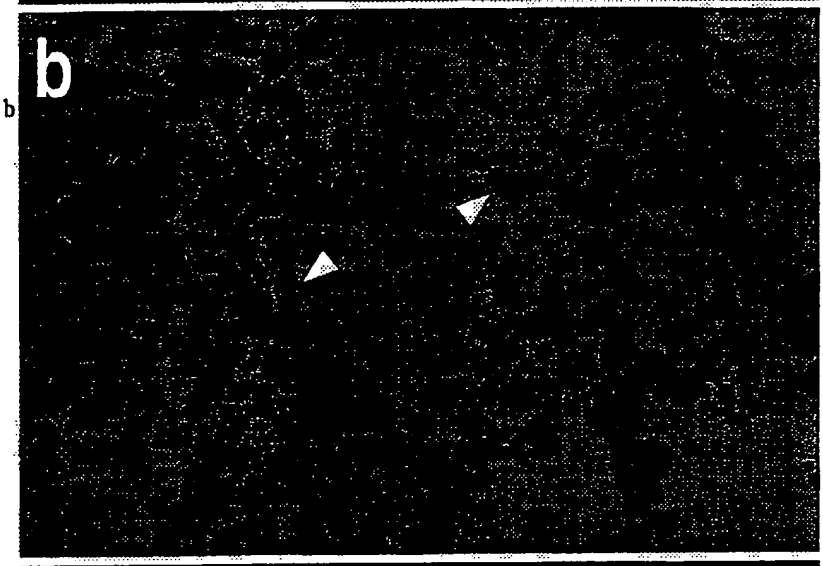
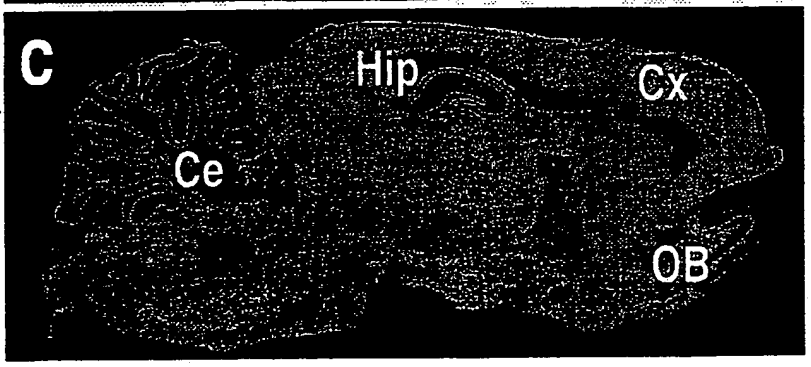


Fig.8 c



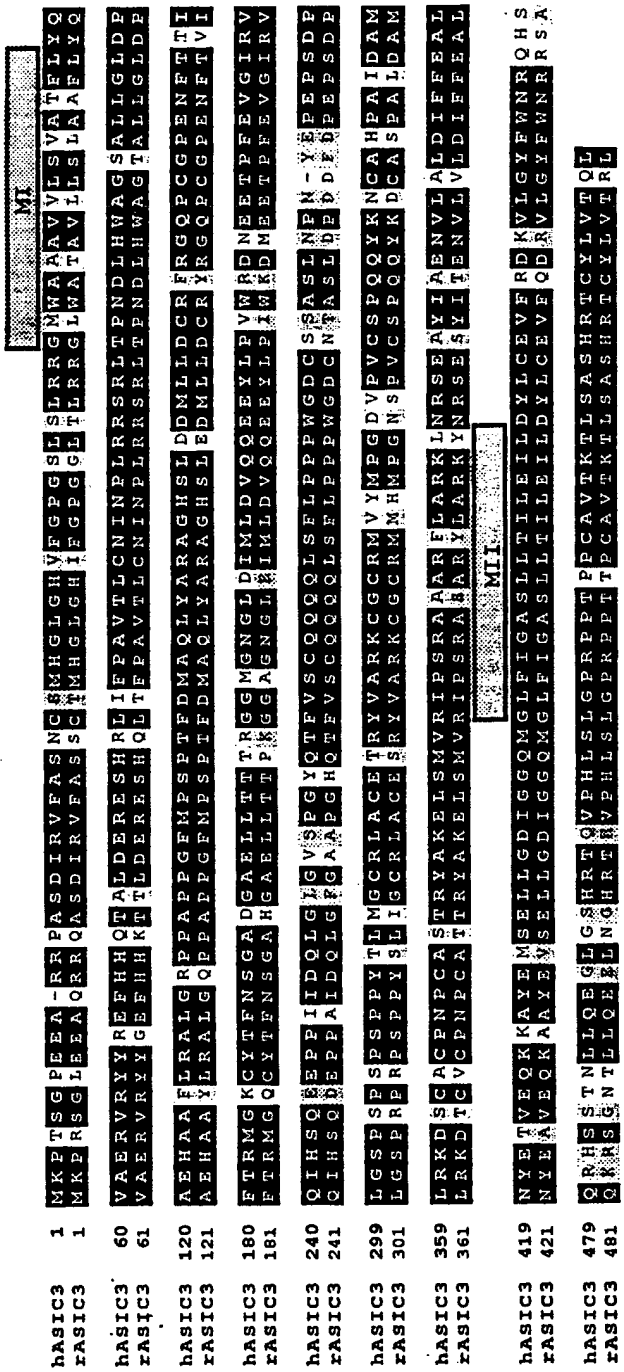

M1
M1
hASIC3 1 MKPTSGPEEA-RRPASDIRVFASNCSMHGLGHVFGPGSLSLRRGMWAAAVVLSVATFLYQ
rASIC3 1 MKPRSGLEEAQRRQASDIRVFASSCTMHGLGHIFGPGGLTLRRGLWATAVTLSLAAELYQ
hASIC3 60 VAERVRYYREFHHQTALDERESHRLIFPAVTLCNINPLRRSRLTPNDLHWAGSALLGLDP
rASIC3 61 VAERVRYYGEFHHKTTLDERESHQLTFPAVTLCNINPLRRSRLTPNDLHWAGTALLGLDP
hASIC3 120 AEHAAFLRALGRPPAPPGFMPSPTFDMAQLYARAGHSLDDMLLDCRFRGQPCGPENFTTI
rASIC3 121 AEHAAYLRALGQPPAPPGFMPSPTFDMAQLYARAGHSLEDMLLDCRYRGQPCGPENFTVI
hASIC3 180 FTRMGKCYTFNSGADGAELLTTTRGGMGNGLDIMLDVQQEEYLPVWRDNEETPFEVGIRV
rASIC3 181 ETRMGQCYTFNSGAHGAELLTTPKGGAGNGLEIMLDVQQEEYLPIWKDMEETPFEVGIRV
hASIC3 240 QIHSQEPPIDQLGLGVSPGYQTFVSCQQQQLSFLPPPWGDCSSASLNPN-YEPEPSDP
rASIC3 241 QIHSQDEPPAIDQLGLGAAPGHQTFVSCQQQQLSFLPPPWGDCNSASLDPDDEDPEPSDP
hASIC3 299 LGSPSPSPPYTLMGCRLACETRYVARKCGCRMVYMPGDVPVCSPQQYKNCAHPAIDAM
rASIC3 301 LGSPRPRPSYSLIGCRLACESRYVARKCGCRMMHMPGNSPVCSPQQYKDCASPALDAM
hASIC3 359 LRKDSCACPNPCASTRYAKELSMVRIPSRAARFLARKLNRSEAYIAENVLALDIFFEAL
rASIC3 361 LRKDTCVCPNPCATTRYAKELSMVRIPSRAARYLARKYNRSESYITENVLVLDIFFEAL
hASIC3 419 NYETVEQKKAYEMSELLGDIGGQMGLFIGASLLTILEILDYLCEVFFRDKVLGYFWNRQHS
rASIC3 421 NYEAVEQKAAYEYSELLGDIGGQMGLFIGASLLTILELLDYLCEVFQDRVLGYFWNRRSA
hASIC3 479 QRHSSTNLLQEGLGSHRTQVPHLSLGPRPPTPPCAVTKTLSASHRTCYLVTQL
rASIC3 481 QRRSLGNTLLQEELNQHRTVPHLSLGPRPPTTPCAVTKTLSASHRTCYLVTRL

Figure 9

Fig.9 continued

Human ASIC3 sequence

```
      10                  30                  50
      .                   .                   .
ACGACGGGGTTCTGGCCATGAAGCCCACCTCAGGCCCAGAGGAGGCCCGGCGGCCAGCCT
                M  K  P  T  S  G  P  E  E  A  R  R  P  A  S

      70                  90                  110
      .                   .                   .
CGGACATCCGCGTGTTCGCCAGCAACTGCTCGATGCACGGGCTGGGCCACGTCTTCGGGC
  D  I  R  V  F  A  S  N  C  S  M  H  G  L  G  H  V  F  G  P

      130                 150                 170
      .                   .                   .
CAGGCAGCCTGAGCCTGCGCCGGGGGATGTGGGCAGCGGCCGTGGTCCTGTCAGTGGCCA
  G  S  L  S  L  R  R  G  M  W  A  A  A  V  V  L  S  V  A  T

      190                 210                 230
      .                   .                   .
CCTTCCTCTACCAGGTGGCTGAGAGGGTGCGCTACTACAGGGAGTTCCACCACCAGACTG
  F  L  Y  Q  V  A  E  R  V  R  Y  Y  R  E  F  H  H  Q  T  A

      250                 270                 290
      .                   .                   .
CCCTGGATGAGCGAGAAAGCCACCGGCTCATCTTCCCGGCTGTCACCCTGTGCAACATCA
  L  D  E  R  E  S  H  R  L  I  F  P  A  V  T  L  C  N  I  N

      310                 330                 350
      .                   .                   .
ACCCACTGCGCCGCTCGCGCCTAACGCCCAACGACCTGCACTGGGCTGGGTCTGCGCTGC
  P  L  R  R  S  R  L  T  P  N  D  L  H  W  A  G  S  A  L  L

      370                 390                 410
      .                   .                   .
TGGGCCTGGATCCCGCAGAGCACGCCGCCTTCCTGCGCGCCCTGGGCCGGCCCCCTGCAC
  G  L  D  P  A  E  H  A  A  F  L  R  A  L  G  R  P  P  A  P
```

Fig.9 continued

```
      430                 450                 470
       .                   .                   .                   .
CGCCCGGCTTCATGCCCAGTCCCACCTTTGACATGGCGCAACTCTATGCCCGTGCTGGGC
  P  G  F  M  P  S  P  T  F  D  M  A  Q  L  Y  A  R  A  G  H

      490                 510                 530
       .                   .                   .                   .
ACTCCCTGGATGACATGCTGCTGGACTGTCGCTTCCGTGGCCAACCTTGTGGGCCTGAGA
  S  L  D  D  M  L  L  D  C  R  F  R  G  Q  P  C  G  P  E  N

      550                 570                 590
       .                   .                   .                   .
ACTTCACCACGATCTTCACCCGGATGGGAAAGTGCTACACATTTAACTCTGGCGCTGATG
  F  T  T  I  F  T  R  M  G  K  C  Y  T  F  N  S  G  A  D  G

      610                 630                 650
       .                   .                   .                   .
GGGCAGAGCTGCTCACCACTACTAGGGGTGGCATGGGCAATGGGCTGGACATCATGCTGG
  A  E  L  L  T  T  T  R  G  G  M  G  N  G  L  D  I  M  L  D

      670                 690                 710
       .                   .                   .                   .
ACGTGCAGCAGGAGGAATATCTACCTGTGTGGAGGGACAATGAGGAGACCCCGTTTGAGG
  V  Q  Q  E  E  Y  L  P  V  W  R  D  N  E  E  T  P  F  E  V

      730                 750                 770
       .                   .                   .                   .
TGGGGATCCGAGTGCAGATCCACAGCCAGGAGGAGCCGCCCATCATCGATCAGCTGGGCT
  G  I  R  V  Q  I  H  S  Q  E  E  P  P  I  I  D  Q  L  G  L

      790                 810                 830
       .                   .                   .                   .
TGGGGGTGTCCCCGGGCTACCAGACCTTTGTTTCTTGCCAGCAGCAGCAGCTGAGCTTCC
  G  V  S  P  G  Y  Q  T  F  V  S  C  Q  Q  Q  Q  L  S  F  L

      850                 870                 890
       .                   .                   .                   .
TGCCACCGCCCTGGGGCGATTGCAGTTCAGCATCTCTGAACCCCAACTATGAGCCAGAGC
```

Fig.9 continued

```
P  P  P  W  G  D  C  S  S  A  S  L  N  P  N  Y  E  P  E  P

      910                 930                 950
CCTCTGATCCCCTAGGCTCCCCCAGCCCCAGCCCCAGCCCTCCCTATACCCTTATGGGGT
S  D  P  L  G  S  P  S  P  S  P  S  P  P  Y  T  L  M  G  C

      970                 990                 1010
GTCGCCTGGCCTGCGAAACCCGCTACGTGGCTCGGAAGTGCGGCTGCCGAATGGTGTACA
R  L  A  C  E  T  R  Y  V  A  R  K  C  G  C  R  M  V  Y  M

      1030                1050                1070
TGCCAGGCGACGTGCCAGTGTGCAGCCCCCAGCAGTACAAGAACTGTGCCCACCCGGCCA
P  G  D  V  P  V  C  S  P  Q  Q  Y  K  N  C  A  H  P  A  I

      1090                1110                1130
TAGATGCCATGCTTCGCAAGGACTCGTGCGCCTGCCCCAACCCGTGCGCCAGCACGCGCT
D  A  M  L  R  K  D  S  C  A  C  P  N  P  C  A  S  T  R  Y

      1150                1170                1190
ACGCCAAGGAGCTCTCCATGGTGCGGATCCCGAGCCGCGCCGCCGCGCGCTTCCTGGCCC
A  K  E  L  S  M  V  R  I  P  S  R  A  A  A  R  F  L  A  R

      1210                1230                1250
GGAAGCTCAACCGCAGCGAGGCCTACATCGCGGAGAACGTGCTGGCCCTGGACATCTTCT
K  L  N  R  S  E  A  Y  I  A  E  N  V  L  A  L  D  I  F  F

      1270                1290                1310
TTGAGGCCCTCAACTATGAGACCGTGGAGCAGAAGAAGGCCTATGAGATGTCAGAGCTGC
E  A  L  N  Y  E  T  V  E  Q  K  K  A  Y  E  M  S  E  L  L
```

Fig.9 continued

```
    1330                1350                1370
      .         .         .         .         .         .
TTGGTGACATTGGGGGCCAGATGGGGCTGTTCATCGGGGCCAGCCTGCTCACCATCCTCG
  G  D  I  G  G  Q  M  G  L  F  I  G  A  S  L  L  T  I  L  E

    1390                1410                1430
      .         .         .         .         .         .
AGATCCTAGACTACCTCTGTGAGGTGTTCCGAGACAAGGTCCTGGGATATTTCTGGAACC
  I  L  D  Y  L  C  E  V  F  R  D  K  V  L  G  Y  F  W  N  R
```

Fig.10 a

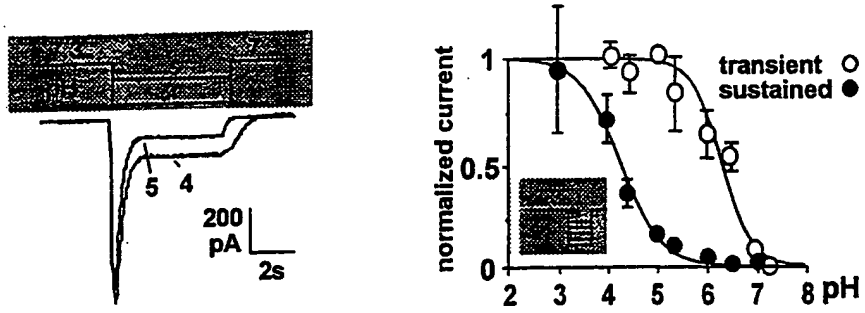


Fig.10 b

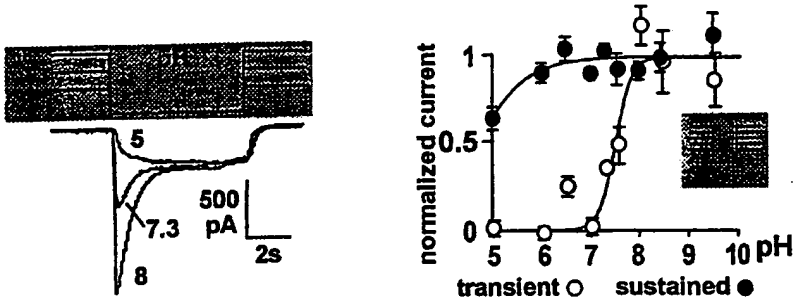


Fig.10 c

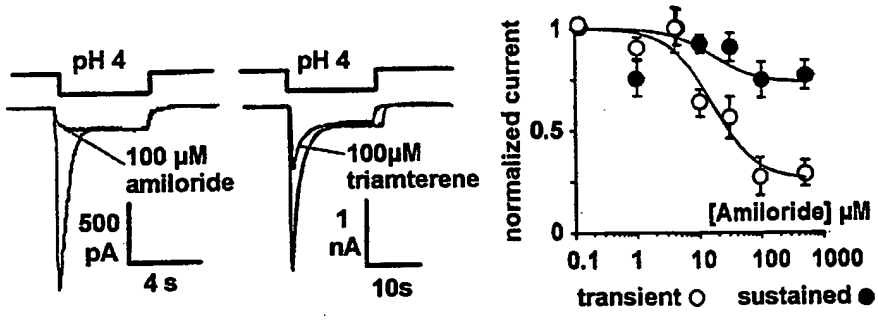


Fig.11 a

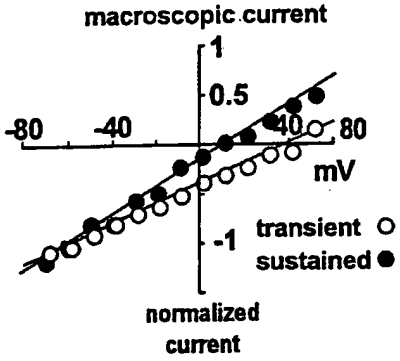


Fig.11 b

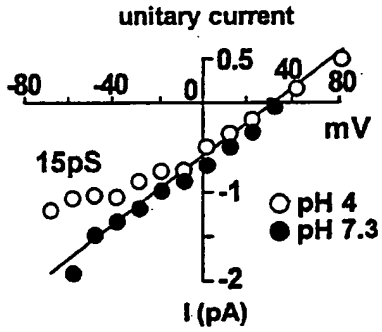


Fig.11 c

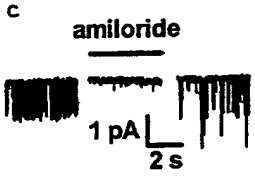
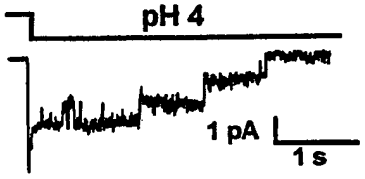


Fig.11 d



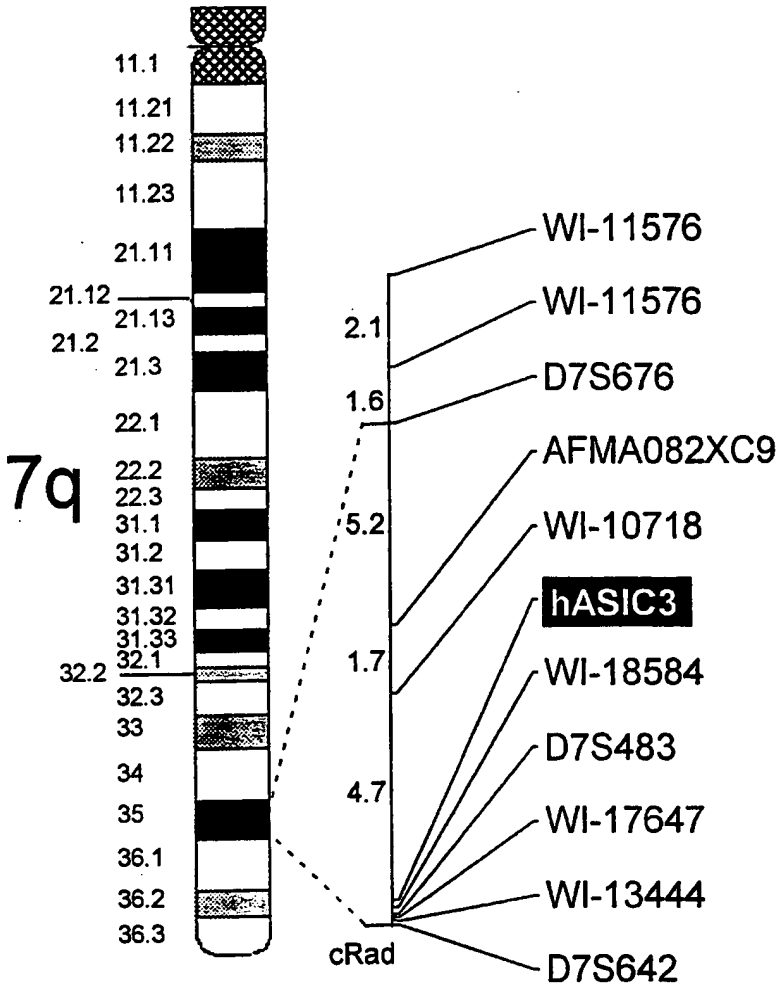

7q
11.1
11.21
11.22
11.23
21.11
21.12
21.13
21.2
21.3
22.1
22.2
22.3
31.1
31.2
31.31
31.32
31.33
32.1
32.2
32.3
33
34
35
36.1
36.2
36.3
2.1
1.6
5.2
1.7
4.7
cRad
WI-11576
WI-11576
D7S676
AFMA082XC9
WI-10718
hASIC3
WI-18584
D7S483
WI-17647
WI-13444
D7S642

Figure 12

LISTE DE SÉQUENCES.

NOMBRE DE SÉQUENCES : 8

INFORMATION CONCERNANT LA SEQ ID NO:1 :

i) CARACTRERISTIQUE DE LA SEQUENCE :
A) LONGUEUR : 3562 paires de base
B) TYPE : acide nucléique
C) NOMBRE DE BRINS : double
D) CONFIGURATION : linéaire

ii) TYPE DE MOLECULE : ADN
vi) ORIGINE : rat

ix) CARACTERISTIQUE
A) NOM/CLE : ASIC
B) LOCALISATION : 123 .. 1700

xi) DESCRIPTION DE LA SEQUENCE : SEQ ID NO:1 :

```
CACACACACA CACACACACA CACACACACA CACACACACA CACACAGAAC CTGCGCCTGT      60

GCCTGTGCCT GTGCCTGTGC CTGTTTGAGA GCTGGAGACA CAGAAGGATC CCCTTGGCAA     120

GG ATG GAA TTG AAG ACC GAG GAG GAG GAG GTG GGT GGT GTC CAG CCG         167
   Met Glu Leu Lys Thr Glu Glu Glu Glu Val Gly Gly Val Gln Pro
   1               5                  10                  15

GTG AGC ATC CAG GCT TTC GCC AGC AGC TCC ACG CTG CAT GGT CTT GCC       215
Val Ser Ile Gln Ala Phe Ala Ser Ser Ser Thr Leu His Gly Leu Ala
                20                  25                  30

CAC ATC TTC TCC TAT GAG CGG CTG TCT CTG AAG CGG GCA CTG TGG GCC       263
His Ile Phe Ser Tyr Glu Arg Leu Ser Leu Lys Arg Ala Leu Trp Ala
            35                  40                  45

CTG TGC TTC CTG GGT TCG CTG GCC GTC CTG CTG TGT GTG TGC ACT GAG       311
Leu Cys Phe Leu Gly Ser Leu Ala Val Leu Leu Cys Val Cys Thr Glu
        50                  55                  60

CGT GTG CAG TAC TAC TTC TGC TAT CAC CAC GTC ACC AAG CTT GAC GAA       359
Arg Val Gln Tyr Tyr Phe Cys Tyr His His Val Thr Lys Leu Asp Glu
    65                  70                  75

GTG GCT GCC TCC CAG CTC ACC TTC CCT GCT GTC ACA CTG TGC AAT CTC       407
Val Ala Ala Ser Gln Leu Thr Phe Pro Ala Val Thr Leu Cys Asn Leu
80                  85                  90                  95

AAT GAG TTC CGC TTT AGC CAA GTC TCC AAG AAT GAC CTG TAC CAT GCT       455
Asn Glu Phe Arg Phe Ser Gln Val Ser Lys Asn Asp Leu Tyr His Ala
                100                 105                 110

GGG GAG CTG CTG GCC CTG CTC AAC AAC AGG TAT GAG ATC CCG GAC ACA       503
Gly Glu Leu Leu Ala Leu Leu Asn Asn Arg Tyr Glu Ile Pro Asp Thr
            115                 120                 125
```

```
CAG ATG GCT GAT GAA AAG CAG CTA GAG ATA TTG CAG GAC AAG GCC AAC      551
Gln Met Ala Asp Glu Lys Gln Leu Glu Ile Leu Gln Asp Lys Ala Asn
        130                 135                 140

TTC CGG AGC TTC AAG CCC AAG CCC TTC AAC ATG CGT GAA TTC TAC GAC      599
Phe Arg Ser Phe Lys Pro Lys Pro Phe Asn Met Arg Glu Phe Tyr Asp
    145                 150                 155

AGA GCG GGG CAC GAT ATT CGA GAC ATG CTG CTC TCG TGC CAC TTC CGT      647
Arg Ala Gly His Asp Ile Arg Asp Met Leu Leu Ser Cys His Phe Arg
160                 165                 170                 175

GGG GAG GCC TGC AGC GCT GAA GAT TTC AAA GTG GTC TTC ACT CGG TAT      695
Gly Glu Ala Cys Ser Ala Glu Asp Phe Lys Val Val Phe Thr Arg Tyr
                180                 185                 190

GGG AAG TGT TAC ACA TTC AAC TCG GGC CAA GAT GGG CGG CCA CGG CTG      743
Gly Lys Cys Tyr Thr Phe Asn Ser Gly Gln Asp Gly Arg Pro Arg Leu
            195                 200                 205

AAG ACC ATG AAA GGT GGG ACT GGC AAT GGC CTG GAG ATC ATG CTG GAC      791
Lys Thr Met Lys Gly Gly Thr Gly Asn Gly Leu Glu Ile Met Leu Asp
        210                 215                 220

ATT CAG CAA GAT GAA TAT TTG CCT GTG TGG GGA GAG ACC GAC GAG ACA      839
Ile Gln Gln Asp Glu Tyr Leu Pro Val Trp Gly Glu Thr Asp Glu Thr
    225                 230                 235

TCC TTC GAA GCA GGC ATC AAA GTG CAG ATC CAC AGT CAG GAT GAA CCC      887
Ser Phe Glu Ala Gly Ile Lys Val Gln Ile His Ser Gln Asp Glu Pro
240                 245                 250                 255

CCT TTC ATC GAC CAG CTG GGC TTT GGT GTG GCT CCA GGT TTC CAG ACG      935
Pro Phe Ile Asp Gln Leu Gly Phe Gly Val Ala Pro Gly Phe Gln Thr
                260                 265                 270

TTT GTG TCT TGC CAG GAG CAG AGG CTC ATC TAC CTG CCC TCA CCC TGG      983
Phe Val Ser Cys Gln Glu Gln Arg Leu Ile Tyr Leu Pro Ser Pro Trp
            275                 280                 285

GGC ACC TGC AAT GCT GTT ACC ATG GAC TCG GAT TTC TTC GAC TCC TAC     1031
Gly Thr Cys Asn Ala Val Thr Met Asp Ser Asp Phe Phe Asp Ser Tyr
        290                 295                 300

AGC ATC ACT GCC TGC CGG ATT GAT TGC GAG ACG CGT TAC CTG GTG GAG     1079
Ser Ile Thr Ala Cys Arg Ile Asp Cys Glu Thr Arg Tyr Leu Val Glu
    305                 310                 315

AAC TGC AAC TGC CGT ATG GTG CAC ATG CCA GGG GAC GCC CCA TAC TGC     1127
Asn Cys Asn Cys Arg Met Val His Met Pro Gly Asp Ala Pro Tyr Cys
320                 325                 330                 335

ACT CCA GAG CAG TAC AAG GAG TGT GCA GAT CCT GCC CTG GAC TTC CTA     1175
Thr Pro Glu Gln Tyr Lys Glu Cys Ala Asp Pro Ala Leu Asp Phe Leu
                340                 345                 350

GTG GAG AAA GAC CAG GAA TAC TGC GTG TGT GAG ATG CCT TGC AAC CTG     1223
Val Glu Lys Asp Gln Glu Tyr Cys Val Cys Glu Met Pro Cys Asn Leu
            355                 360                 365
```

```
ACC CGC TAC GGC AAG GAG CTG TCC ATG GTC AAG ATC CCA AGC AAA GCC      1271
Thr Arg Tyr Gly Lys Glu Leu Ser Met Val Lys Ile Pro Ser Lys Ala
        370                 375                 380

TCC GCC AAG TAC CTG GCC AAG AAG TTC AAC AAA TCG GAG CAG TAC ATA      1319
Ser Ala Lys Tyr Leu Ala Lys Lys Phe Asn Lys Ser Glu Gln Tyr Ile
    385                 390                 395

GGG GAG AAC ATT CTG GTG CTG GAC ATT TTC TTT GAA GTC CTC AAC TAT      1367
Gly Glu Asn Ile Leu Val Leu Asp Ile Phe Phe Glu Val Leu Asn Tyr
400                 405                 410                 415

GAG ACC ATC GAG CAG AAA AAG GCC TAT GAG ATC GCA GGG CTG TTG GGT      1415
Glu Thr Ile Glu Gln Lys Lys Ala Tyr Glu Ile Ala Gly Leu Leu Gly
                420                 425                 430

GAC ATC GGG GGC CAG ATG GGG TTG TTC ATC GGT GCC AGC ATC CTC ACC      1463
Asp Ile Gly Gly Gln Met Gly Leu Phe Ile Gly Ala Ser Ile Leu Thr
            435                 440                 445

GTG CTG GAA CTC TTT GAC TAT GCC TAC GAG GTC ATT AAG CAC AGG CTG      1511
Val Leu Glu Leu Phe Asp Tyr Ala Tyr Glu Val Ile Lys His Arg Leu
        450                 455                 460

TGC AGA CGT GGA AAG TGC CAG AAG GAG GCT AAG AGG AGC AGC GCA GAC      1559
Cys Arg Arg Gly Lys Cys Gln Lys Glu Ala Lys Arg Ser Ser Ala Asp
    465                 470                 475

AAG GGC GTG GCG CTC AGC CTG GAT GAC GTC AAA AGA CAC AAT CCC TGC      1607
Lys Gly Val Ala Leu Ser Leu Asp Asp Val Lys Arg His Asn Pro Cys
480                 485                 490                 495

GAG AGC CTC CGA GGA CAT CCT GCC GGG ATG ACG TAC GCT GCC AAC ATC      1655
Glu Ser Leu Arg Gly His Pro Ala Gly Met Thr Tyr Ala Ala Asn Ile
                500                 505                 510

CTA CCT CAC CAT CCC GCT CGA GGC ACG TTT GAG GAC TTT ACC TGC TAA      1703
Leu Pro His His Pro Ala Arg Gly Thr Phe Glu Asp Phe Thr Cys  *
            515                 520                 526

GCCCTCGCAG GCCGCTGTAC CAAAGGCCTA GGTGGGGAGG GCTGGGGGAG CAAGGGGCCC    1763

CCAACTGCCC CCAGCTACCC TGTGGACTTA ACTGCATTCC TGGTCAGTGG TTCCCTCTTG    1823

TCTGTGGTGA GAAAGGAGTC TTGACCATAG AGTCCTCTCC CAGCCTCTAT CCCATCTTTT    1943

TATTTTAATT TAATCACATT TGCTCTGTAA TATTGCTTGA GGCTGGGGAT CGTGATTTCC    2003

CCCCAGTTCT TTTATTGTTG AGAATAGTTT TCTCTATTCT GGGTTTTCTG TTATTTCAAA    2063
TGAATCTGCA AATTGCTCTT CCCATCTCTA TGAAGAATTG CGTTGGAATT TTGATGGGGA    2123

TTGTATTGAA TCTGTAGATT GCCTTTGGTA AGATGGCCAT TTTTACTATG TTAATCCTGC    2183

CAATTCATGA GCAAGGGAGA TCTTTCTATC TCTGAAATCT ACTTCAGTTT CTTTCTTCAG    2243

AGACTTGAAG TTCTTGTCAT AAAAATCTTT TTGGTTAGAG CCACACCAAG GTATTTTATA    2303

TTGTTTGTGA CTATTGTGAA TGGTGTCATT TCCCTAATTT CCTTCTCAGC CTACTTATCC    2363
```

TTTGAGTAGA GGAAGGCTTC TGATTTGTTT GGGTTAATTT TATACCCAGC TGCTTTGCTA 2423

AAGTTCTTTA TCAGGTTTAG GTGTTCTCTG GTGGAACTTT TGGGGTCACG TAAGAATACT 2483

ATTATATCAT CTGCAAATAG TGATATTTCA CTTCTTCCTT TCCAATTTCT ATCCCTCTGG 2543

GGACTTTTTG TTGTCTAATT GCTCTGGCTA GGACTTCAAA TTCTATATTG AATAGATAGG 2603

GAGAGAGTGG GCAGCCTTGT CTAGTTCCTG GTTTTCGTGG GATCGCTTCA AATTTCTCTC 2663

CATTTAGTTT GATATTGGCT ACTGGTTTGC TGTATATGGC TTTTACTGTA CTTAGGTATG 2723

GGCCTTGAAT TCCTGATATT TCCAAGACTT TTAACATGAA GGGGTTTTGA AATTTGCCAA 2783

ATGCTTTCTC AGCATCTAAT GAGATGATCA TGTGCCCTCC CCCCACCTTG AGTTTGTTTA 2843

TATAGTGGGT TACATGAAAG GATCATTTCT AATAGTCCAC AAGTCTGCCA AATCTTGCTG 2903

ATTGTGACTC ATTTCCATAG CAGGCTCTAT AACTTCTCTA ACAGATTGCA TTAAACTCTG 2963

CTTGGGGAAG GCATTACCTC TTGGTTGAAG CAATGTTGTA GTTTCTATGC CTGCTGAGTA 3023

AATAGCCTCA AGTCCAAGTA CTTGCCCAGA CTAATGATCA AACGTATCCA GGAGTTCCAT 3083

ACCAGAGATG TACTCTTCTC TCCTTTGAAG TACATTGCTG GAAGAGTAAT TGTGTTTGCT 3143

AGAGATACTC CTTCGAACTG CAAAAGAAAT CTCTTGGCTA AGCATATAAT CAAGCCTCAG 3203

GTTTTCTTTT TATTAAATAG CTGCTTGTAA GAAAGTGGAC ACTAAGCATA TACCTCAAAG 3263

GGAGACAGAA TGACTCTGTG CCTTCACTGA TGGAAGTCTG GGTTACAAAT TACATCAGAA 3323

GAACCTATCA TAGTGAAACA TCTCATTCCC CTGGTATAAT CCCTTCTAGA AATACACTTG 3383

TGACTCTGAA ATGTTATAAT CGTGACAACT AGGCTGTTAC AGATACACCA AGTTAAATTT 3443

GATAGAGAAA CCAGGCTTGG AGCCTCATGT CCATAGGGCA AGAGGAAGAT GCTGAGTGTT 3503

TAAGGTTGGT TTGAGCGAAG AACAATACCT TGTGTCACAA AAATGAAAGG AAAAAAGAAA 3563

AAAGGAAAGA AGGAAAGAAA GAGAGAGAAA GAAAAAGAAA GAAAGAAAAA AAAAAAAAA 3562

INFORMATION CONCERNANT LA SEQ ID NO:2 :

i) CARACTRERISTIQUE DE LA SEQUENCE :
A) LONGUEUR : 1620 paires de base
B) TYPE : acide nucléique
C) NOMBRE DE BRINS : double
D) CONFIGURATION : linéaire

ii) TYPE DE MOLECULE : ADN
vi) ORIGINE : homme

ix) CARACTERISTIQUE
A) NOM/CLE : ASIC
B) LOCALISATION : 1 .. 1542

xi) DESCRIPTION DE LA SEQUENCE : SEQ ID NO:2 :

```
CCG GTG AGC ATC CAG GCC TTC GCC AGC AGC TCC ACA CTG CAC GGC ATG        48
Pro Val Ser Ile Gln Ala Phe Ala Ser Ser Ser Thr Leu His Gly Met
1               5                   10                  15

GCC CAC ATC TTC TCC TAC GAG CGG CTG TCT CTG AAG CGG GCA CTG TGG        96
Ala His Ile Phe Ser Tyr Glu Arg Leu Ser Leu Lys Arg Ala Leu Trp
            20                  25                  30

GCC CTG TGC TTC CTG GGC TCG CTG GCT GTG CTG CTG TGT GTG TGC ACG       144
Ala Leu Cys Phe Leu Gly Ser Leu Ala Val Leu Leu Cys Val Cys Thr
        35                  40                  45

GAG CGT GTG CAG TAC TAC TTC CAC TAc CAC CAT GTC ACC AAG CTC GAC       192
Glu Arg Val Gln Tyr Tyr Phe His Tyr His His Val Thr Lys Leu Asp
    50                  55                  60

GAG GTG GCT GCC TCT CAG CTT ACC TTC CCT GCT GTC ACG CTG TGC AAC       240
Glu Val Ala Ala Ser Gln Leu Thr Phe Pro Ala Val Thr Leu Cys Asn
65                  70                  75                  80

CTC AAC GAG TTC CGC TTT AGC CAA GTC TCC AAG AAT GAC CTG TAT CAT       288
Leu Asn Glu Phe Arg Phe Ser Gln Val Ser Lys Asn Asp Leu Tyr His
                85                  90                  95

GCT GGG GAG CTG CTG GCC CTG CTC AAC AAC AGG TAT GAG ATA CCA GAC       336
Ala Gly Glu Leu Leu Ala Leu Leu Asn Asn Arg Tyr Glu Ile Pro Asp
            100                 105                 110

ACA CAG ATG GCA GAT GAA AAG CAG CTG GAG ATA CTG CAG GAC AAA GCC       384
Thr Gln Met Ala Asp Glu Lys Gln Leu Glu Ile Leu Gln Asp Lys Ala
        115                 120                 125

AAC TTC CGC AGC TTC AAA CCC AAA CCC TTC AAC ATG CGT GAG TTC TAC       432
Asn Phe Arg Ser Phe Lys Pro Lys Pro Phe Asn Met Arg Glu Phe Tyr
    130                 135                 140

GAC CGA GCT GGG CAC GAC ATT CGA GAC ATG CTG CTC TCC TGC CAC TTC       480
Asp Arg Ala Gly His Asp Ile Arg Asp Met Leu Leu Ser Cys His Phe
145                 150                 155                 160
```

```
CGG GGG GAG GTC TGC AGC GCT GAA GAC TTC AAG GTG GTC TTC ACA CGC      528
Arg Gly Glu Val Cys Ser Ala Glu Asp Phe Lys Val Val Phe Thr Arg
                165                 170                 175

TAT GGA AAG TGC TAC ACG TTC AAC TCG GGC CGA AAT GGG CGG CCG CGG      576
Tyr Gly Lys Cys Tyr Thr Phe Asn Ser Gly Arg Asn Gly Arg Pro Arg
            180                 185                 190

CTG AAG ACC ATG AAG GGT GGG ACG GGC AAT GGG CTG GAA ATC ATG CTG      624
Leu Lys Thr Met Lys Gly Gly Thr Gly Asn Gly Leu Glu Ile Met Leu
        195                 200                 205

GAC ATC CAG CAG GAC GAG TAC CTG CCT GTG TGG GGG GAG ACT GAC GAG      672
Asp Ile Gln Gln Asp Glu Tyr Leu Pro Val Trp Gly Glu Thr Asp Glu
    210                 215                 220

ACG TCT TTC GAA GCA GGC ATC AAA GTG CAG ATC CAT AGT CAG GAT GAA      720
Thr Ser Phe Glu Ala Gly Ile Lys Val Gln Ile His Ser Gln Asp Glu
225                 230                 235                 240

CCT CCT TTC ATC GAC CAG CTG GGC TTT GGC GTG GCC CCA GGC TTC CAG      768
Pro Pro Phe Ile Asp Gln Leu Gly Phe Gly Val Ala Pro Gly Phe Gln
                245                 250                 255

ACC TTT GTG GCC TGC CAG GAG CAG CGG CTC ATA TAC CTG CCC CCA CCC      816
Thr Phe Val Ala Cys Gln Glu Gln Arg Leu Ile Tyr Leu Pro Pro Pro
            260                 265                 270

TGG GGC ACC TGC AAA GCT GTT ACC ATG GAC TCG GAT TTG GAT TTC TTC      864
Trp Gly Thr Cys Lys Ala Val Thr Met Asp Ser Asp Leu Asp Phe Phe
        275                 280                 285

GAC TCC TAC AGC ATC ACT GCC TGC CGC ATC GAC TGT GAG ACG CGC TAC      912
Asp Ser Tyr Ser Ile Thr Ala Cys Arg Ile Asp Cys Glu Thr Arg Tyr
    290                 295                 300

CTG GTG GAG AAC TGC AAC TGC CGC ATG GTG CAC ATG CCA GGG GAT GCC      960
Leu Val Glu Asn Cys Asn Cys Arg Met Val His Met Pro Gly Asp Ala
305                 310                 315                 320

CCA TAC TGT ACT CCA GAG CAG TAC AAG GAG TGT GCA GAT CCT GCT CTG     1008
Pro Tyr Cys Thr Pro Glu Gln Tyr Lys Glu Cys Ala Asp Pro Ala Leu
                325                 330                 335

GAC TTC CTG GTG GAG AAG GAC CAG GAG TAC TGC GTG TGT GAA ATG CCT     1056
Asp Phe Leu Val Glu Lys Asp Gln Glu Tyr Cys Val Cys Glu Met Pro
            340                 345                 350

TGC AAC CTG ACC CGC TAT GGC AAA GAG CTG TCC ATG GTC AAG ATC CCC     1104
Cys Asn Leu Thr Arg Tyr Gly Lys Glu Leu Ser Met Val Lys Ile Pro
        355                 360                 365

AGC AAA GCC TCA GCC AAG TAC CTG GCC AAG AAG TTC AAC AAA TCT GAG     1152
Ser Lys Ala Ser Ala Lys Tyr Leu Ala Lys Lys Phe Asn Lys Ser Glu
    370                 375                 380

CAA TAC ATA GGG GAG AAC ATC CTG GTG CTG GAC ATT TTC TTT GAA GTC     1200
Gln Tyr Ile Gly Glu Asn Ile Leu Val Leu Asp Ile Phe Phe Glu Val
385                 390                 395                 400
```

```
CTC AAC TAT GAG ACC ATT GAA CAG AAG AAG GCC TAT GAG ATT GCA GGG      1248
Leu Asn Tyr Glu Thr Ile Glu Gln Lys Lys Ala Tyr Glu Ile Ala Gly
                405                 410                 415

CTC CTG GGT GAC ATC GGG GGC CAG ATG GGG CTG TTC ATC GGG GCC AGC      1296
Leu Leu Gly Asp Ile Gly Gly Gln Met Gly Leu Phe Ile Gly Ala Ser
            420                 425                 430

ATC CTC ACG GTG CTG GAG CTC TTT GAC TAC GCC TAC GGG GTC ATT AAG      1344
Ile Leu Thr Val Leu Glu Leu Phe Asp Tyr Ala Tyr Gly Val Ile Lys
        435                 440                 445

CAC AAG CTG TGC CGA CGA GGA AAA TGC CAG AAG GAG GCC AAA AGG AGC      1392
His Lys Leu Cys Arg Arg Gly Lys Cys Gln Lys Glu Ala Lys Arg Ser
    450                 455                 460

AGT GCG GAC AAG GGC GTG GCC CTC AGC CTG GAC GAC GTC AAA AGA CAC      1440
Ser Ala Asp Lys Gly Val Ala Leu Ser Leu Asp Asp Val Lys Arg His
465                 470                 475                 480

AAC CCG TGC GAG AGC CTT CGG GGC CAC CCT GCC GGG ATG ACA TAC GCT      1488
Asn Pro Cys Glu Ser Leu Arg Gly His Pro Ala Gly Met Thr Tyr Ala
                485                 490                 495

GCC AAC ATC GTA CCT CAC CAT CCG GCC CGA GGC ACG TTC GAG GAC TTT      1536
Ala Asn Ile Val Pro His His Pro Ala Arg Gly Thr Phe Glu Asp Phe
            500                 505                 510

ACC TGC TGA GCCCCGCAGG CCGCCGAACC AAAGACCTAG ATGGGGAGGA CTAGGAGAGC   1595
Thr Cys *
    514

GAGGGGGCCC CCAGCTGCCT CCTAA                                         1620
```

INFORMATION CONCERNANT LA SEQ ID NO:3 :

i) CARACTRERISTIQUE DE LA SEQUENCE :
A) LONGUEUR : 1666 paires de base
B) TYPE : acide nucléique
C) NOMBRE DE BRINS : double
D) CONFIGURATION : linéaire

ii) TYPE DE MOLECULE : ADN
vi) ORIGINE : homme

ix) CARACTERISTIQUE
A) NOM/CLE : MDEG
B) LOCALISATION : 127 .. 1663

xi) DESCRIPTION DE LA SEQUENCE : SEQ ID NO:3 :

```
TCTGGCGCGA TGCTTACCTT GCGTTCTCTC CCCTGAACGT CAAGGTTTAA GCAGAGCCCG      60

AGGACTGGGA GCTCTTCTCT GAAATTCGAT CAACCTGAAG CCAGTTGCGG AACTGCACGG     120

GGTCCCG ATG GAC CTC AAG GAA AGC CCC AGT GAG GGC AGC CTG CAA CCT        169
        Met Asp Leu Lys Glu Ser Pro Ser Glu Gly Ser Leu Gln Pro
        1               5                   10

TCT AGC ATC CAG ATC TTT GCC AAC ACC TCC ACC CTC CAT GGC ATC CGC        217
Ser Ser Ile Gln Ile Phe Ala Asn Thr Ser Thr Leu His Gly Ile Arg
15                  20                  25                  30

CAC ATC TTC GTG TAT GGG CCG CTG ACC ATC CGG CGT GTG CTG TGG GCA        265
His Ile Phe Val Tyr Gly Pro Leu Thr Ile Arg Arg Val Leu Trp Ala
                35                  40                  45

GTG GCC TTC GTG GGC TCT CTG GGC CTG CTG CTG GTG GAG AGC TCT GAG        313
Val Ala Phe Val Gly Ser Leu Gly Leu Leu Leu Val Glu Ser Ser Glu
            50                  55                  60

AGG GTG TCC TAC TAC TTC TCC TAC CAG CAT GTC ACT AAG GTG GAC GAA        361
Arg Val Ser Tyr Tyr Phe Ser Tyr Gln His Val Thr Lys Val Asp Glu
        65                  70                  75

GTG GTG GCT CAA AGC CTG GTC TTC CCA GCT GTG ACC CTC TGT AAC CTC        409
Val Val Ala Gln Ser Leu Val Phe Pro Ala Val Thr Leu Cys Asn Leu
    80                  85                  90

AAT GGC TTC CGG TTC TCC AGG CTC ACC ACC AAC GAC CTG TAC CAT GCT        457
Asn Gly Phe Arg Phe Ser Arg Leu Thr Thr Asn Asp Leu Tyr His Ala
95                  100                 105                 110

GGG GAG CTG CTG GCC CTG CTG GAT GTC AAC CTG CAG ATC CCG GAC CCC        505
Gly Glu Leu Leu Ala Leu Leu Asp Val Asn Leu Gln Ile Pro Asp Pro
                115                 120                 125

CAT CTG GCT GAC CCC TCC GTG CTG GAG GCC CTG CGG CAG AAG GCC AAC        553
His Leu Ala Asp Pro Ser Val Leu Glu Ala Leu Arg Gln Lys Ala Asn
            130                 135                 140
```

```
TTC AAG CAC TAC AAA CCC AAG CAG TTC AGC ATG CTG GAG TTC CTG CAC        601
Phe Lys His Tyr Lys Pro Lys Gln Phe Ser Met Leu Glu Phe Leu His
        145                 150                 155

CGT GTG GGC CAT GAC CTG AAG GAT ATG ATG CTC TAC TGC AAG TTC AAA        649
Arg Val Gly His Asp Leu Lys Asp Met Met Leu Tyr Cys Lys Phe Lys
    160                 165                 170

GGG CAG GAG TGC GGC CAC CAA GAC TTC ACC ACA GTG TTT ACA AAA TAT        697
Gly Gln Glu Cys Gly His Gln Asp Phe Thr Thr Val Phe Thr Lys Tyr
175                 180                 185                 190

GGG AAG TGT TAC ATG TTT AAC TCA GGC GAG GAT GGC AAA CCT CTG CTC        745
Gly Lys Cys Tyr Met Phe Asn Ser Gly Glu Asp Gly Lys Pro Leu Leu
                195                 200                 205

ACC ACG GTC AAG GGG GGG ACA GGC AAC GGG CTG GAG ATC ATG CTG GAC        793
Thr Thr Val Lys Gly Gly Thr Gly Asn Gly Leu Glu Ile Met Leu Asp
            210                 215                 220

ATT CAG CAG GAT GAG TAC CTG CCC ATC TGG GGA GAG ACA GAG GAA ACG        841
Ile Gln Gln Asp Glu Tyr Leu Pro Ile Trp Gly Glu Thr Glu Glu Thr
        225                 230                 235

ACA TTT GAA GCA GGA GTG AAA GTT CAG ATC CAC AGT CAG TCT GAG CCA        889
Thr Phe Glu Ala Gly Val Lys Val Gln Ile His Ser Gln Ser Glu Pro
    240                 245                 250

CCT TTC ATC CAA GAG CTG GGC TTT GGG GTG GCT CCA GGG TTC CAG ACC        937
Pro Phe Ile Gln Glu Leu Gly Phe Gly Val Ala Pro Gly Phe Gln Thr
255                 260                 265                 270

TTT GTG GCC ACA CAG GAG CAG AGG CTC ACA TAC CTG CCC CCA CCG TGG        985
Phe Val Ala Thr Gln Glu Gln Arg Leu Thr Tyr Leu Pro Pro Pro Trp
                275                 280                 285

GGT GAG TGC CGA TCC TCA GAG ATG GGC CTC GAC TTT TTT CCT GTT TAC       1033
Gly Glu Cys Arg Ser Ser Glu Met Gly Leu Asp Phe Phe Pro Val Tyr
            290                 295                 300

AGC ATC ACC GCC TGT AGG ATT GAC TGT GAG ACC CGC TAC ATT GTG GAA       1081
Ser Ile Thr Ala Cys Arg Ile Asp Cys Glu Thr Arg Tyr Ile Val Glu
        305                 310                 315

AAC TGC AAC TGC CGC ATG GTT CAC ATG CCA GGG GAT GCC CCT TTT TGT       1129
Asn Cys Asn Cys Arg Met Val His Met Pro Gly Asp Ala Pro Phe Cys
    320                 325                 330

ACC CCT GAG CAG CAC AAG GAG TGT GCA GAG CCT GCC CTA GGT CTG TTG       1177
Thr Pro Glu Gln His Lys Glu Cys Ala Glu Pro Ala Leu Gly Leu Leu
335                 340                 345                 350

GCG GAA AAG GAC AGC AAT TAC TGT CTC TGC AGG ACA CCC TGC AAC CTA       1225
Ala Glu Lys Asp Ser Asn Tyr Cys Leu Cys Arg Thr Pro Cys Asn Leu
                355                 360                 365

ACC CGC TAC AAC AAA GAG CTC TCC ATG GTG AAG ATC CCC AGC AAG ACA       1273
Thr Arg Tyr Asn Lys Glu Leu Ser Met Val Lys Ile Pro Ser Lys Thr
            370                 375                 380
```

```
TCA GCC AAG TAC CTT GAG AAG AAA TTT AAC AAA TCA GAA AAA TAT ATC      1321
Ser Ala Lys Tyr Leu Glu Lys Lys Phe Asn Lys Ser Glu Lys Tyr Ile
        385                 390                 395

TCA GAG AAC ATC CTT GTT CTG GAT ATA TTT TTT GAA GCT CTC AAT TAT      1369
Ser Glu Asn Ile Leu Val Leu Asp Ile Phe Phe Glu Ala Leu Asn Tyr
    400                 405                 410

GAG ACA ATT GAA CAG AAG AAG GCG TAT GAA GTT GCT GCC TTA CTT GGT      1417
Glu Thr Ile Glu Gln Lys Lys Ala Tyr Glu Val Ala Ala Leu Leu Gly
415                 420                 425                 430

GAT ATT GGT GGT CAG ATG GGA TTG TTC ATT GGT GCT AGT ATC CTT ACA      1465
Asp Ile Gly Gly Gln Met Gly Leu Phe Ile Gly Ala Ser Ile Leu Thr
                435                 440                 445

ATA CTA GAG CTC TTT GAT TAT ATT TAT GAG CTG ATC AAA GAG AAG CTA      1513
Ile Leu Glu Leu Phe Asp Tyr Ile Tyr Glu Leu Ile Lys Glu Lys Leu
            450                 455                 460

TTA GAC CTG CTT GGC AAA GAG GAG GAT GAA GGG AGC CAC GAT GAG AAT      1561
Leu Asp Leu Leu Gly Lys Glu Glu Asp Glu Gly Ser His Asp Glu Asn
        465                 470                 475

GTG AGT ACT TGT GAC ACA ATG CCA AAC CAC TCT GAA ACC ATC AGT CAC      1609
Val Ser Thr Cys Asp Thr Met Pro Asn His Ser Glu Thr Ile Ser His
    480                 485                 490

ACT GTG AAC GTG CCC CTG CAG ACG ACC CTG GGG ACC CTG GAA GAA ATA      1657
Thr Val Asn Val Pro Leu Gln Thr Thr Leu Gly Thr Leu Glu Glu Ile
495                 500                 505                 510

GCC TGC TGA                                                          1666
Ala Cys  *
    512
```

INFORMATION CONCERNANT LA SEQ ID NO:4 :

i) CARACTRERISTIQUE DE LA SEQUENCE :
A) LONGUEUR : 3647 paires de base
B) TYPE : acide nucléique
C) NOMBRE DE BRINS : double
D) CONFIGURATION : linéaire

ii) TYPE DE MOLECULE : ADN
vi) ORIGINE : rat

ix) CARACTERISTIQUE
A) NOM/CLE : ASIC1B
B) LOCALISATION : 109 .. 1785

xi) DESCRIPTION DE LA SEQUENCE : SEQ ID NO:4 :

```
CTGCCACAGA GGCTCTGGTG AGGAAGGACA GACAGCTGGA CCGGCGCAGA CCTAGCCGAA      60

GTCCAACCTC CGTCCCTTCT GGTGGCTTCT TCCTGTCTCC TGAACAAG ATG CCC ATC      117
                                                        Met Pro Ile
                                                        1       3

CAG ATC TTT TGT TCT GTG TCA TTC TCC TCT GGA GAG GAG GCC CCG GGA        165
Gln Ile Phe Cys Ser Val Ser Phe Ser Ser Gly Glu Glu Ala Pro Gly
    5                   10                  15

TCC ATG GCA GAT ATC TGG GGT CCC CAC CAC CAC CGG CAG CAG CAG GAC        213
Ser Met Ala Asp Ile Trp Gly Pro His His His Arg Gln Gln Gln Asp
20                  25                  30                  35

AGC TCA GAA TCG GAA GAA GAG GAA GAG AAG GAA ATG GAG GCA GGG TCG        261
Ser Ser Glu Ser Glu Glu Glu Glu Glu Lys Glu Met Glu Ala Gly Ser
                40                  45                  50

GAG TTG GAT GAG GGT GAT GAC TCA CCT AGG GAC TTG GTG GCC TTC GCC        309
Glu Leu Asp Glu Gly Asp Asp Ser Pro Arg Asp Leu Val Ala Phe Ala
            55                  60                  65

AAC AGC TGT ACC TTC CAT GGT GCC AGC CAT GTG TTT GTG GAA GGG GGC        357
Asn Ser Cys Thr Phe His Gly Ala Ser His Val Phe Val Glu Gly Gly
        70                  75                  80

CCA GGG CCA AGG CAG GCC TTA TGG GCA GTG GCC TTT GTC ATA GCA CTG        405
Pro Gly Pro Arg Gln Ala Leu Trp Ala Val Ala Phe Val Ile Ala Leu
    85                  90                  95

GGT GCC TTC CTG TGC CAG GTA GGG GAC CGC GTT GCT TAT TAC CTC AGC        453
Gly Ala Phe Leu Cys Gln Val Gly Asp Arg Val Ala Tyr Tyr Leu Ser
100                 105                 110                 115

TAC CCA CAC GTG ACT TTG CTA GAC GAA GTG GCC ACC ACG GAG CTG GTC        501
Tyr Pro His Val Thr Leu Leu Asp Glu Val Ala Thr Thr Glu Leu Val
                120                 125                 130
```

```
TTC CCA GCT GTC ACC TTC TGC AAC ACC AAT GCC GTG CGG TTG TCC CAG      549
Phe Pro Ala Val Thr Phe Cys Asn Thr Asn Ala Val Arg Leu Ser Gln
            135                 140                 145

CTC AGC TAC CCT GAC TTG CTC TAC CTG GCC CCC ATG CTA GGA CTG GAT      597
Leu Ser Tyr Pro Asp Leu Leu Tyr Leu Ala Pro Met Leu Gly Leu Asp
        150                 155                 160

GAG AGT GAT GAC CCC GGG GTG CCC CTT GCT CCT CCT GGC CCA GAG GCT      645
Glu Ser Asp Asp Pro Gly Val Pro Leu Ala Pro Pro Gly Pro Glu Ala
    165                 170                 175

TTC TCC GGG GAG CCT TTT AAC CTC CAT CGT TTC TAT AAT CGC TCT TGC      693
Phe Ser Gly Glu Pro Phe Asn Leu His Arg Phe Tyr Asn Arg Ser Cys
180                 185                 190                 195

CAC CGG CTG GAG GAC ATG CTG CTC TAT TGT TCC TAC TGT GGG GGC CCC      741
His Arg Leu Glu Asp Met Leu Leu Tyr Cys Ser Tyr Cys Gly Gly Pro
                200                 205                 210

TGT GGT CCC CAC AAC TTC TCA GTG GTC TTC ACT CGG TAT GGG AAG TGT      789
Cys Gly Pro His Asn Phe Ser Val Val Phe Thr Arg Tyr Gly Lys Cys
            215                 220                 225

TAC ACA TTC AAC TCG GGC CAA GAT GGG CGG CCA CGG CTG AAG ACC ATG      837
Tyr Thr Phe Asn Ser Gly Gln Asp Gly Arg Pro Arg Leu Lys Thr Met
        230                 235                 240

AAA GGT GGG ACT GGC AAT GGC CTG GAG ATC ATG CTG GAC ATT CAG CAA      885
Lys Gly Gly Thr Gly Asn Gly Leu Glu Ile Met Leu Asp Ile Gln Gln
    245                 250                 255

GAT GAA TAT TTG CCT GTG TGG GGA GAG ACC GAC GAG ACA TCC TTC GAA      933
Asp Glu Tyr Leu Pro Val Trp Gly Glu Thr Asp Glu Thr Ser Phe Glu
260                 265                 270                 275

GCA GGC ATC AAA GTG CAG ATC CAC AGT CAG GAT GAA CCC CCT TTC ATC      981
Ala Gly Ile Lys Val Gln Ile His Ser Gln Asp Glu Pro Pro Phe Ile
                280                 285                 290

GAC CAG CTG GGC TTT GGT GTG GCT CCA GGT TTC CAG ACG TTT GTG TCT     1029
Asp Gln Leu Gly Phe Gly Val Ala Pro Gly Phe Gln Thr Phe Val Ser
            295                 300                 305

TGC CAG GAG CAG AGG CTC ATC TAC CTG CCC TCA CCC TGG GGC ACC TGC     1077
Cys Gln Glu Gln Arg Leu Ile Tyr Leu Pro Ser Pro Trp Gly Thr Cys
        310                 315                 320

AAT GCT GTT ACC ATG GAC TCG GAT TTC TTC GAC TCC TAC AGC ATC ACT     1125
Asn Ala Val Thr Met Asp Ser Asp Phe Phe Asp Ser Tyr Ser Ile Thr
    325                 330                 335

GCC TGC CGG ATT GAT TGC GAG ACG CGT TAC CTG GTG GAG AAC TGC AAC     1173
Ala Cys Arg Ile Asp Cys Glu Thr Arg Tyr Leu Val Glu Asn Cys Asn
340                 345                 350                 355
```

```
TGC CGT ATG GTG CAC ATG CCA GGG GAC GCC CCA TAC TGC ACT CCA GAG      1221
Cys Arg Met Val His Met Pro Gly Asp Ala Pro Tyr Cys Thr Pro Glu
                360                 365                 370

CAG TAC AAG GAG TGT GCA GAT CCT GCC CTG GAC TTC CTA GTG GAG AAA      1269
Gln Tyr Lys Glu Cys Ala Asp Pro Ala Leu Asp Phe Leu Val Glu Lys
            375                 380                 385

GAC CAG GAA TAC TGC GTG TGT GAG ATG CCT TGC AAC CTG ACC CGC TAC      1317
Asp Gln Glu Tyr Cys Val Cys Glu Met Pro Cys Asn Leu Thr Arg Tyr
        390                 395                 400

GGC AAG GAG CTG TCC ATG GTC AAG ATC CCA AGC AAA GCC TCC GCC AAG      1365
Gly Lys Glu Leu Ser Met Val Lys Ile Pro Ser Lys Ala Ser Ala Lys
    405                 410                 415

TAC CTG GCC AAG AAG TTC AAC AAA TCG GAG CAG TAC ATA GGG GAG AAC      1413
Tyr Leu Ala Lys Lys Phe Asn Lys Ser Glu Gln Tyr Ile Gly Glu Asn
420                 425                 430                 435

ATT CTG GTG CTG GAC ATT TTC TTT GAA GTC CTC AAC TAT GAG ACC ATC      1461
Ile Leu Val Leu Asp Ile Phe Phe Glu Val Leu Asn Tyr Glu Thr Ile
                440                 445                 450

GAG CAG AAA AAG GCC TAT GAG ATC GCA GGG CTG TTG GGT GAC ATC GGG      1509
Glu Gln Lys Lys Ala Tyr Glu Ile Ala Gly Leu Leu Gly Asp Ile Gly
            455                 460                 465

GGC CAG ATG GGG TTG TTC ATC GGT GCC AGC ATC CTC ACC GTG CTG GAA      1557
Gly Gln Met Gly Leu Phe Ile Gly Ala Ser Ile Leu Thr Val Leu Glu
        470                 475                 480

CTC TTT GAC TAT GCC TAC GAG GTC ATT AAG CAC AGG CTG TGC AGA CGT      1605
Leu Phe Asp Tyr Ala Tyr Glu Val Ile Lys His Arg Leu Cys Arg Arg
    485                 490                 495

GGA AAG TGC CAG AAG GAG GCT AAG AGG AGC AGC GCA GAC AAG GGC GTG      1653
Gly Lys Cys Gln Lys Glu Ala Lys Arg Ser Ser Ala Asp Lys Gly Val
500                 505                 510                 515

GCG CTC AGC CTG GAT GAC GTC AAA AGA CAC AAT CCC TGC GAG AGC CTC      1701
Ala Leu Ser Leu Asp Asp Val Lys Arg His Asn Pro Cys Glu Ser Leu
                520                 525                 530

CGA GGA CAT CCT GCC GGG ATG ACG TAC GCT GCC AAC ATC CTA CCT CAC      1749
Arg Gly His Pro Ala Gly Met Thr Tyr Ala Ala Asn Ile Leu Pro His
            535                 540                 545

CAT CCC GCT CGA GGC ACG TTT GAG GAC TTT ACC TGC TAA GCCCTCGCAG       1798
His Pro Ala Arg Gly Thr Phe Glu Asp Phe Thr Cys  *
        550                 55                  559

GCCGCTGTAC CAAAGGCCTA GGTGGGGAGG GCTGGGGGAG CAAGGGGCCC CCAACTGCCC    1858

CCAGCTACCC TGTGGACTTA ACTGCATTCC TGGTCAGTGG TTCCCTCTTG TCTGTGGTGA    1918

GAAAGGAGTC TTGACCATAG AGTCCTCTCC CAGCCTCTAT CCCATCTTTT TATTTTAATT    1978

TAATCACATT TGCTCTGTAA TATTGCTTGA GGCTGGGGAT CGTGATTTCC CCCCAGTTCT    2038
```

TTTATTGTTG AGAATAGTTT TCTCTATTCT GGGTTTTCTG TTATTTCAAA TGAATCTGCA 2098

AATTGCTCTT CCCATCTCTA TGAAGAATTG CGTTGGAATT TTGATGGGGA TTGTATTGAA 2158

TCTGTAGATT GCCTTTGGTA AGATGGCCAT TTTTACTATG TTAATCCTGC CAATTCATGA 2218

GCAAGGGAGA TCTTTCTATC TCTGAAATCT ACTTCAGTTT CTTTCTTCAG AGACTTGAAG 2278

TTCTTGTCAT AAAAATCTTT TTGGTTAGAG CCACACCAAG GTATTTTATA TTGTTTGTGA 2338

CTATTGTGAA TGGTGTCATT TCCCTAATTT CCTTCTCAGC CTACTTATCC TTTGAGTAGA 2398

GGAAGGCTTC TGATTTGTTT GGGTTAATTT TATACCCAGC TGCTTTGCTA AAGTTCTTTA 2458

TCAGGTTTAG GTGTTCTCTG GTGGAACTTT TGGGGTCACG TAAGAATACT ATTATATCAT 2518

CTGCAAATAG TGATATTTCA CTTCTTCCTT TCCAATTTCT ATCCCTCTGG GGACTTTTTG 2578

TTGTCTAATT GCTCTGGCTA GGACTTCAAA TTCTATATTG AATAGATAGG GAGAGAGTGG 2638

GCAGCCTTGT CTAGTTCCTG GTTTTCGTGG GATCGCTTCA AATTTCTCTC CATTTAGTTT 2698

GATATTGGCT ACTGGTTTGC TGTATATGGC TTTTACTGTA CTTAGGTATG GGCCTTGAAT 2758

TCCTGATATT TCCAAGACTT TTAACATGAA GGGGTTTTGA AATTTGCCAA ATGCTTTCTC 2818

AGCATCTAAT GAGATGATCA TGTGCCCTCC CCCCACCTTG AGTTTGTTTA TATAGTGGGT 2878

TACATGAAAG GATCATTTCT AATAGTCCAC AAGTCTGCCA AATCTTGCTG ATTGTGACTC 2938

ATTTCCATAG CAGGCTCTAT AACTTCTCTA ACAGATTGCA TTAAACTCTG CTTGGGGAAG 2998

GCATTACCTC TTGGTTGAAG CAATGTTGTA GTTTCTATGC CTGCTGAGTA AATAGCCTCA 3058

AGTCCAAGTA CTTGCCCAGA CTAATGATCA AACGTATCCA GGAGTTCCAT ACCAGAGATG 3118

TACTCTTCTC TCCTTTGAAG TACATTGCTG GAAGAGTAAT TGTGTTTGCT AGAGATACTC 3178

CTTCGAACTG CAAAAGAAAT CTCTTGGCTA AGCATATAAT CAAGCCTCAG GTTTTCTTTT 3238

TATTAAATAG CTGCTTGTAA GAAAGTGGAC ACTAAGCATA TACCTCAAAG GGAGACAGAA 3298

TGACTCTGTG CCTTCACTGA TGGAAGTCTG GGTTACAAAT TACATCAGAA GAACCTATCA 3358

TAGTGAAACA TCTCATTCCC CTGGTATAAT CCCTTCTAGA AATACACTTG TGACTCTGAA 3418

ATGTTATAAT CGTGACAACT AGGCTGTTAC AGATACACCA AGTTAAATTT GATAGAGAAA 3478

CCAGGCTTGG AGCCTCATGT CCATAGGGCA AGAGGAAGAT GCTGAGTGTT TAAGGTTGGT 3538

TTGAGCGAAG AACAATACCT TGTGTCACAA AAATGAAAGG AAAAAAGAAA AAAGGAAAGA 3598

AGGAAAGAAA GAGAGAGAAA GAAAAAGAAA GAAAGAAAAA AAAAAAAAA 3647

INFORMATION CONCERNANT LA SEQ ID NO:5 :

i) CARACTRERISTIQUE DE LA SEQUENCE :
A) LONGUEUR 1602 paires de base
B) TYPE : acide nucléique
C) NOMBRE DE BRINS : double
D) CONFIGURATION : linéaire

ii) TYPE DE MOLECULE : ADN
vi) ORIGINE : rat

ix) CARACTERISTIQUE
A) NOM/CLE : DRASIC
B) LOCALISATION : 1 .. 1602

xi) DESCRIPTION DE LA SEQUENCE : SEQ ID NO:5 :

```
ATG AAA CCT CGC TCC GGA CTG GAG GAG GCC CAG CGG CGA CAG GCC TCA        48
Met Lys Pro Arg Ser Gly Leu Glu Glu Ala Gln Arg Arg Gln Ala Ser
1               5                  10                  15

GAC ATC CGG GTG TTT GCC AGC AGC TGC ACA ATG CAT GGT CTG GGC CAC        96
Asp Ile Arg Val Phe Ala Ser Ser Cys Thr Met His Gly Leu Gly His
            20                  25                  30

ATC TTT GGC CCT GGA GGC CTG ACC CTG CGC CGA GGG CTG TGG GCC ACA       144
Ile Phe Gly Pro Gly Gly Leu Thr Leu Arg Arg Gly Leu Trp Ala Thr
        35                  40                  45

GCT GTG CTC CTG TCG CTG GCG GCC TTC CTC TAC CAG GTG GCT GAG CGG       192
Ala Val Leu Leu Ser Leu Ala Ala Phe Leu Tyr Gln Val Ala Glu Arg
    50                  55                  60

GTT CGC TAC TAT GGG GAG TTC CAC CAT AAG ACC ACC CTG GAT GAG CGT       240
Val Arg Tyr Tyr Gly Glu Phe His His Lys Thr Thr Leu Asp Glu Arg
65                  70                  75                  80

GAG AGC CAC CAG CTC ACC TTC CCA GCT GTG ACT CTG TGT AAT ATC AAC       288
Glu Ser His Gln Leu Thr Phe Pro Ala Val Thr Leu Cys Asn Ile Asn
                85                  90                  95

CCA CTG CGC CGC TCA CGC CTC ACA CCC AAT GAC TTG CAC TGG GCT GGA       336
Pro Leu Arg Arg Ser Arg Leu Thr Pro Asn Asp Leu His Trp Ala Gly
            100                 105                 110

ACA GCG CTG CTG GGC CTG GAC CCT GCT GAA CAT GCT GCC TAC CTT CGT       384
Thr Ala Leu Leu Gly Leu Asp Pro Ala Glu His Ala Ala Tyr Leu Arg
        115                 120                 125

GCA CTG GGC CAG CCC CCC GCA CCA CCT GGC TTC ATG CCC AGT CCG ACC       432
Ala Leu Gly Gln Pro Pro Ala Pro Pro Gly Phe Met Pro Ser Pro Thr
    130                 135                 140

TTT GAC ATG GCA CAA CTC TAC GCC AGA GCC GGC CAC TCC CTT GAG GAC       480
Phe Asp Met Ala Gln Leu Tyr Ala Arg Ala Gly His Ser Leu Glu Asp
145                 150                 155                 160
```

```
ATG TTG TTG GAT TGC CGA TAC CGT GGC CAG CCC TGT GGG CCT GAG AAC      528
Met Leu Leu Asp Cys Arg Tyr Arg Gly Gln Pro Cys Gly Pro Glu Asn
                165                 170                 175

TTC ACA GTG ATC TTT ACT CGA ATG GGG CAA TGC TAC ACC TTC AAC TCT      576
Phe Thr Val Ile Phe Thr Arg Met Gly Gln Cys Tyr Thr Phe Asn Ser
            180                 185                 190

GGT GCC CAC GGT GCA GAG CTG CTC ACC ACT CCA AAG GGT GGT GCT GGC      624
Gly Ala His Gly Ala Glu Leu Leu Thr Thr Pro Lys Gly Gly Ala Gly
        195                 200                 205

AAC GGA CTG GAG ATT ATG CTA GAT GTA CAG CAA GAG GAG TAT CTG CCC      672
Asn Gly Leu Glu Ile Met Leu Asp Val Gln Gln Glu Glu Tyr Leu Pro
    210                 215                 220

ATC TGG AAG GAC ATG GAA GAG ACC CCG TTT GAG GTG GGG ATC CGA GTG      720
Ile Trp Lys Asp Met Glu Glu Thr Pro Phe Glu Val Gly Ile Arg Val
225                 230                 235                 240

CAG ATT CAC AGC CAG GAT GAG CCC CCT GCC ATT GAC CAG CTG GGC TTC      768
Gln Ile His Ser Gln Asp Glu Pro Pro Ala Ile Asp Gln Leu Gly Phe
                245                 250                 255

GGG GCA GCC CCA GGC CAT CAG ACT TTT GTG TCC TGT CAG CAG CAG CAA      816
Gly Ala Ala Pro Gly His Gln Thr Phe Val Ser Cys Gln Gln Gln Gln
            260                 265                 270

CTG AGT TTC CTG CCA CCA CCC TGG GGT GAC TGC AAT ACC GCA TCT TTG      864
Leu Ser Phe Leu Pro Pro Pro Trp Gly Asp Cys Asn Thr Ala Ser Leu
        275                 280                 285

GAT CCC GAC GAC TTT GAT CCA GAG CCC TCT GAT CCC TTG GGT TCC CCC      912
Asp Pro Asp Asp Phe Asp Pro Glu Pro Ser Asp Pro Leu Gly Ser Pro
    290                 295                 300

AGA CCC AGA CCC AGC CCT CCT TAT AGT TTA ATA GGT TGT CGC CTG GCC      960
Arg Pro Arg Pro Ser Pro Pro Tyr Ser Leu Ile Gly Cys Arg Leu Ala
305                 310                 315                 320

TGT GAG TCT CGC TAT GTG GCT CGG AAG TGT GGC TGT CGA ATG ATG CAT     1008
Cys Glu Ser Arg Tyr Val Ala Arg Lys Cys Gly Cys Arg Met Met His
                325                 330                 335

ATG CCT GGA AAC TCC CCA GTG TGC AGC CCC CAG CAG TAC AAG GAC TGC     1056
Met Pro Gly Asn Ser Pro Val Cys Ser Pro Gln Gln Tyr Lys Asp Cys
            340                 345                 350

GCC AGC CCA GCT CTG GAC GCT ATG CTG CGA AAG GAC ACG TGT GTC TGC     1104
Ala Ser Pro Ala Leu Asp Ala Met Leu Arg Lys Asp Thr Cys Val Cys
        355                 360                 365

CCC AAC CCG TGC GCT ACT ACA CGC TAT GCC AAG GAG CTC TCC ATG GTG     1152
Pro Asn Pro Cys Ala Thr Thr Arg Tyr Ala Lys Glu Leu Ser Met Val
    370                 375                 380

CGG ATT CCC AGC CGC GCG TCA GCT CGC TAC CTG GCC CGG AAA TAC AAC     1200
Arg Ile Pro Ser Arg Ala Ser Ala Arg Tyr Leu Ala Arg Lys Tyr Asn
385                 390                 395                 400
```

```
CGC AGC GAG TCC TAC ATT ACG GAG AAT GTA CTG GTT CTG GAT ATC TTC      1248
Arg Ser Glu Ser Tyr Ile Thr Glu Asn Val Leu Val Leu Asp Ile Phe
                405                 410                 415

TTT GAG GCC CTC AAC TAT GAA GCG GTG GAA CAA AAG GCG GCC TAT GAA      1296
Phe Glu Ala Leu Asn Tyr Glu Ala Val Glu Gln Lys Ala Ala Tyr Glu
            420                 425                 430

GTG TCG GAG CTG CTG GGA GAC ATT GGG GGA CAG ATG GGA CTG TTT ATT      1344
Val Ser Glu Leu Leu Gly Asp Ile Gly Gly Gln Met Gly Leu Phe Ile
        435                 440                 445

GGA GCA AGC CTG CTT ACC ATC CTT GAG ATC CTC GAC TAT CTC TGT GAG      1392
Gly Ala Ser Leu Leu Thr Ile Leu Glu Ile Leu Asp Tyr Leu Cys Glu
    450                 455                 460

GTT TTC CAA GAC AGA GTC CTG GGG TAT TTC TGG AAC AGA AGG AGC GCT      1440
Val Phe Gln Asp Arg Val Leu Gly Tyr Phe Trp Asn Arg Arg Ser Ala
465                 470                 475                 480

CAA AAG CGC TCT GGC AAC ACT CTG CTC CAG GAA GAG TTG AAT GGC CAT      1488
Gln Lys Arg Ser Gly Asn Thr Leu Leu Gln Glu Glu Leu Asn Gly His
                485                 490                 495

CGA ACA CAT GTT CCC CAC CTC AGC CTA GGG CCC AGG CCT CCT ACC ACT      1536
Arg Thr His Val Pro His Leu Ser Leu Gly Pro Arg Pro Pro Thr Thr
            500                 505                 510

CCC TGT GCT GTC ACC AAG ACA CTC TCT GCC TCC CAC CGT ACC TGT TAC      1584
Pro Cys Ala Val Thr Lys Thr Leu Ser Ala Ser His Arg Thr Cys Tyr
        515                 520                 525

CTC GTC ACA AGG CTC TAG                                              1602
Leu Val Thr Arg Leu  *
    530             533
```

INFORMATION CONCERNANT LA SEQ ID NO:6 :

i) CARACTRERISTIQUE DE LA SEQUENCE :
A) LONGUEUR 1948 paires de base
B) TYPE : acide nucléique
C) NOMBRE DE BRINS : double
D) CONFIGURATION : linéaire

ii) TYPE DE MOLECULE : ADN
vi) ORIGINE : rat

ix) CARACTERISTIQUE
A) NOM/CLE : MDEG2
B) LOCALISATION : 16 .. 1707

xi) DESCRIPTION DE LA SEQUENCE : SEQ ID NO:6 :

```
CCTCGGGCTG AATGA ATG AGC CGG AGC GGC GGA GCC CGG CTG CCC GCG ACC        51
                 Met Ser Arg Ser Gly Gly Ala Arg Leu Pro Ala Thr
                 1               5                   10

GCG CTC AGC GGC CCG GGA CGC TTC CGT ATG GCC CGC GAG CAG CCG GCG        99
Ala Leu Ser Gly Pro Gly Arg Phe Arg Met Ala Arg Glu Gln Pro Ala
        15                  20                  25

CCC GTG GCG GTG GCG GCA GCT AGG CAG CCC GGA GGA GAC CGG AGC GGC       147
Pro Val Ala Val Ala Ala Ala Arg Gln Pro Gly Gly Asp Arg Ser Gly
    30                  35                  40

GAT CCG GCG CTG CAG GGG CCA GGG GTC GCC CGC AGG GGG CGG CCG TCC       195
Asp Pro Ala Leu Gln Gly Pro Gly Val Ala Arg Arg Gly Arg Pro Ser
45                  50                  55                  60

CTG AGT CGC ACT AAA TTG CAC GGG CTG CGG CAC ATG TGC GCG GGG CGC       243
Leu Ser Arg Thr Lys Leu His Gly Leu Arg His Met Cys Ala Gly Arg
                65                  70                  75

ACG GCG GCG GGA GGC TCT TTC CAG CGA CGG GCG CTG TGG GTG CTG GCC       291
Thr Ala Ala Gly Gly Ser Phe Gln Arg Arg Ala Leu Trp Val Leu Ala
            80                  85                  90

TTC TGC ACG TCC CTC GGC TTG CTG CTG TCC TGG TCC TCG AAC CGC CTG       339
Phe Cys Thr Ser Leu Gly Leu Leu Leu Ser Trp Ser Ser Asn Arg Leu
        95                  100                 105

CTC TAC TGG CTC AGC TTC CCG TCA CAC ACA CGA GTG CAC CGT GAG TGG       387
Leu Tyr Trp Leu Ser Phe Pro Ser His Thr Arg Val His Arg Glu Trp
    110                 115                 120

AGC CGC CAG CTG CCG TTC CCC GCC GTC ACC GTG TGC AAC AAC AAC CCC       435
Ser Arg Gln Leu Pro Phe Pro Ala Val Thr Val Cys Asn Asn Asn Pro
125                 130                 135                 140

CTG CGC TTC CCG CGC CTC TCC AAG GGG GAC CTC TAC TAC GCG GGC CAC       483
Leu Arg Phe Pro Arg Leu Ser Lys Gly Asp Leu Tyr Tyr Ala Gly His
                145                 150                 155
```

```
TGG CTA GGG CTG CTG CTT CCC AAC CGC ACC GCG CGC CCG CTG GTC AGC      531
Trp Leu Gly Leu Leu Leu Pro Asn Arg Thr Ala Arg Pro Leu Val Ser
            160                 165                 170

GAG CTG CTG CGG GGC GAC GAG CCG CGC CGC CAG TGG TTC CGC AAA CTG      579
Glu Leu Leu Arg Gly Asp Glu Pro Arg Arg Gln Trp Phe Arg Lys Leu
        175                 180                 185

GCC GAC TTC CGC CTC TTC CTG CCG CCG CGC CAC TTC GAG GGC ATC AGC      627
Ala Asp Phe Arg Leu Phe Leu Pro Pro Arg His Phe Glu Gly Ile Ser
    190                 195                 200

GCT GCC TTC ATG GAC CGT TTG GGC CAC CAG CTG GAG GAT ATG CTG CTC      675
Ala Ala Phe Met Asp Arg Leu Gly His Gln Leu Glu Asp Met Leu Leu
205                 210                 215                 220

TCC TGC AAG TAC CGG GGC GAG CTC TGT GGC CCG CAC AAC TTC TCC TCA      723
Ser Cys Lys Tyr Arg Gly Glu Leu Cys Gly Pro His Asn Phe Ser Ser
                225                 230                 235

GTG TTT ACA AAA TAC GGG AAG TGT TAC ATG TTT AAC TCA GGC GAG GAT      771
Val Phe Thr Lys Tyr Gly Lys Cys Tyr Met Phe Asn Ser Gly Glu Asp
            240                 245                 250

GGC AAG CCG CTG CTC ACC ACG GTC AAG GGG GGG ACG GGC AAC GGG CTG      819
Gly Lys Pro Leu Leu Thr Thr Val Lys Gly Gly Thr Gly Asn Gly Leu
        255                 260                 265

GAG ATC ATG CTG GAC ATT CAG CAA GAT GAG TAC CTG CCC ATC TGG GGA      867
Glu Ile Met Leu Asp Ile Gln Gln Asp Glu Tyr Leu Pro Ile Trp Gly
    270                 275                 280

GAG ACA GAG GAA ACA ACG TTT GAA GCA GGA GTG AAG GTT CAG ATC CAC      915
Glu Thr Glu Glu Thr Thr Phe Glu Ala Gly Val Lys Val Gln Ile His
285                 290                 295                 300

AGT CAG TCT GAG CCG CCT TTC ATC CAA GAG CTG GGC TTT GGG GTG GCT      963
Ser Gln Ser Glu Pro Pro Phe Ile Gln Glu Leu Gly Phe Gly Val Ala
                305                 310                 315

CCG GGG TTC CAG ACC TTC GTG GCC ACA CAA GAG CAG AGG CTC ACA TAT     1011
Pro Gly Phe Gln Thr Phe Val Ala Thr Gln Glu Gln Arg Leu Thr Tyr
            320                 325                 330

CTG CCC CCA CCA TGG GGG GAG TGC CGG TCC TCA GAG ATG GGA CTC GAC     1059
Leu Pro Pro Pro Trp Gly Glu Cys Arg Ser Ser Glu Met Gly Leu Asp
        335                 340                 345

TTC TTT CCT GTT TAC AGC ATC ACA GCC TGT CGG ATT GAC TGT GAG ACC     1107
Phe Phe Pro Val Tyr Ser Ile Thr Ala Cys Arg Ile Asp Cys Glu Thr
    350                 355                 360

CGC TAC ATC GTG GAG AAC TGT AAC TGC CGC ATG GTC CAC ATG CCA GGG     1155
Arg Tyr Ile Val Glu Asn Cys Asn Cys Arg Met Val His Met Pro Gly
365                 370                 375                 380
```

```
GAC GCC CCT TTC TGC ACC CCT GAG CAG CAC AAG GAG TGT GCA GAG CCT      1203
Asp Ala Pro Phe Cys Thr Pro Glu Gln His Lys Glu Cys Ala Glu Pro
                385                 390                 395

GCC CTC GGT CTA CTG GCA GAA AAG GAC AGC AAT TAC TGT CTC TGC AGG      1251
Ala Leu Gly Leu Leu Ala Glu Lys Asp Ser Asn Tyr Cys Leu Cys Arg
            400                 405                 410

ACA CCC TGC AAC CTG ACA CGC TAC AAC AAA GAG CTC TCC ATG GTG AAG      1299
Thr Pro Cys Asn Leu Thr Arg Tyr Asn Lys Glu Leu Ser Met Val Lys
        415                 420                 425

ATC CCC AGC AAG ACG TCA GCC AAG TAC TTA GAG AAG AAA TTT AAC AAA      1347
Ile Pro Ser Lys Thr Ser Ala Lys Tyr Leu Glu Lys Lys Phe Asn Lys
    430                 435                 440

TCG GAA AAA TAT ATC TCA GAG AAC ATT CTT GTT CTG GAC ATA TTT TTT      1395
Ser Glu Lys Tyr Ile Ser Glu Asn Ile Leu Val Leu Asp Ile Phe Phe
445                 450                 455                 460

GAG GCG CTC AAT TAC GAA ACA ATT GAA CAG AAG AAG GCG TAT GAA GTT      1443
Glu Ala Leu Asn Tyr Glu Thr Ile Glu Gln Lys Lys Ala Tyr Glu Val
                465                 470                 475

GCT GCC TTA CTT GGT GAC ATC GGT GGT CAG ATG GGA CTG TTC ATT GGT      1491
Ala Ala Leu Leu Gly Asp Ile Gly Gly Gln Met Gly Leu Phe Ile Gly
            480                 485                 490

GCT AGT CTC CTC ACA ATA CTA GAG CTC TTT GAT TAT ATT TAT GAG CTG      1539
Ala Ser Leu Leu Thr Ile Leu Glu Leu Phe Asp Tyr Ile Tyr Glu Leu
        495                 500                 505

ATC AAA GAG AAG CTA TTA GAC CTG CTT GGC AAA GAA GAA GAG GAA GGG      1587
Ile Lys Glu Lys Leu Leu Asp Leu Leu Gly Lys Glu Glu Glu Glu Gly
    510                 515                 520

AGC CAC GAT GAG AAC ATG AGC ACC TGT GAC ACA ATG CCA AAC CAC TCT      1635
Ser His Asp Glu Asn Met Ser Thr Cys Asp Thr Met Pro Asn His Ser
525                 530                 535                 540

GAA ACC ATC AGC CAC ACT GTG AAC GTG CCC CTG CAG ACA GCT TTG GGC      1683
Glu Thr Ile Ser His Thr Val Asn Val Pro Leu Gln Thr Ala Leu Gly
                545                 550                 555

ACC CTG GAG GAG ATT GCC TGC TGA CACCTCTCAG GCAACGCAGC ACCTCCAAAC      1737
Thr Leu Glu Glu Ile Ala Cys  *
            560             563

AGACCTTAAA GGCCCAAGAC CTAGGACAGG AGACAGCAAG CGCAGGTGGG ATCGCCCCTG     1797
ACGACTGAAA GAAGCAGAGC CCCCCATATG CACACATTGC GAACTTCTGC CAAACCTCAC     1857
CTGGCCACAT CTGACATGAA CCGTCCCGGG CCCTGCGTCA TGTCCCTCGC AGGACCGATG     1917
AGTCGCACTC CGGAACTGTC CAAGAACTAA C                                    1948
```

SEQ ID No. 7

ACGACGGGGTTCTGGCCATGAAGCCCACCTCAGGCCCAGAGGAGGCCCGGCGGCCAGCCT
CGGACATCCGCGTGTTCGCCAGCAACTGCTCGATGCACGGGCTGGGCCACGTCTTCGGGC
CAGGCAGCCTGAGCCTGCGCCGGGGGATGTGGGCAGCGGCCGTGGTCCTGTCAGTGGCCA
CCTTCCTCTACCAGGTGGCTGAGAGGGTGCGCTACTACAGGGAGTTCCACCACCAGACTG
CCCTGGATGAGCGAGAAAGCCACCGGCTCATCTTCCCGGCTGTCACCCTGTGCAACATCA
ACCCACTGCGCCGCTCGCGCCTAACGCCCAACGACCTGCACTGGGCTGGGTCTGCGCTGC
TGGGCCTGGATCCCGCAGAGCACGCCGCCTTCCTGCGCGCCCTGGGCCGGCCCCCTGCAC
CGCCCGGCTTCATGCCCAGTCCCACCTTTGACATGGCGCAACTCTATGCCCGTGCTGGGC
ACTCCCTGGATGACATGCTGCTGGACTGTCGCTTCCGTGGCCAACCTTGTGGGCCTGAGA
ACTTCACCACGATCTTCACCCGGATGGGAAAGTGCTACACATTTAACTCTGGCGCTGATG
GGGCAGAGCTGCTCACCACTACTAGGGGTGGCATGGGCAATGGGCTGGACATCATGCTGG
ACGTGCAGCAGGAGGAATATCTACCTGTGTGGAGGGACAATGAGGAGACCCCGTTTGAGG
TGGGGATCCGAGTGCAGATCCACAGCCAGGAGGAGCCGCCCATCATCGATCAGCTGGGCT
TGGGGGTGTCCCCGGGCTACCAGACCTTTGTTTCTTGCCAGCAGCAGCAGCTGAGCTTCC
TGCCACCGCCCTGGGGCGATTGCAGTTCAGCATCTCTGAACCCCAACTATGAGCCAGAGC
CCTCTGATCCCCTAGGCTCCCCCAGCCCCAGCCCCAGCCCTCCCTATACCCTTATGGGGT
GTCGCCTGGCCTGCGAAACCCGCTACGTGGCTCGGAAGTGCGGCTGCCGAATGGTGTACA
TGCCAGGCGACGTGCCAGTGTGCAGCCCCCAGCAGTACAAGAACTGTGCCCACCCGGCCA
TAGATGCCATGCTTCGCAAGGACTCGTGCGCCTGCCCCAACCCGTGCGCCAGCACGCGCT
ACGCCAAGGAGCTCTCCATGGTGCGGATCCCGAGCCGCGCCGCCGCGCGCTTCCTGGCCC
GGAAGCTCAACCGCAGCGAGGCCTACATCGCGGAGAACGTGCTGGCCCTGGACATCTTCT
TTGAGGCCCTCAACTATGAGACCGTGGAGCAGAAGAAGGCCTATGAGATGTCAGAGCTGC
TTGGTGACATTGGGGGCCAGATGGGGCTGTTCATCGGGGCCAGCCTGCTCACCATCCTCG
AGATCCTAGACTACCTCTGTGAGGTGTTCCGAGACAAGGTCCTGGGATATTTCTGGAACC
GACAGCACTCCCAAAGGCACTCCAGCACCAATCTGCTTCAGGAAGGGCTGGGCAGCCATC
GAACCCAAGTTCCCCACCTCAGCCTGGGCCCCAGACCTCCCACCCCTCCCTGTGCCGTCA
CCAAGACTCTCTCCGCCTCCCACCGCACCTGCTACCTTGTCACACAGCTCTAGACCTGCT
GTCTGTGTCCTCGGAGCCCCGCCCTGACATCCTGGACATGCCTAGCCTGCACGTAGCTTT
TCCGTCTTCACCCCAAATAAAGTCCTAATGCATCAAAAAAAAAAAAAAAAAAAAAA

SEQ ID No. 8

M	K	P	T	S	G	P	E	E	A	R	R	P	A	S	D	I	R	V	F	A	S
N	C	S	M	H	G	L	G	H	V	F	G	P	G	S	L	S	L	R	R	G	M
W	A	A	A	V	V	L	S	V	A	T	F	L	Y	Q	V	A	E	R	V	R	Y
Y	R	E	F	H	H	Q	T	A	L	D	E	R	E	S	H	R	L	I	F	P	A
V	T	L	C	N	I	N	P	L	R	R	S	R	L	T	P	N	D	L	H	W	A
G	S	A	L	L	G	L	D	P	A	E	H	A	A	F	L	R	A	L	G	R	P
P	A	P	P	G	F	M	P	S	P	T	F	D	M	A	Q	L	Y	A	R	A	G
H	S	L	D	D	M	L	L	D	C	R	F	R	G	Q	P	C	G	P	E	N	F
T	T	I	F	T	R	M	G	K	C	Y	T	F	N	S	G	A	D	G	A	E	L
L	T	T	T	R	G	G	M	G	N	G	L	D	I	M	L	D	V	Q	Q	E	E
Y	L	P	V	W	R	D	N	E	E	T	P	F	E	V	G	I	R	V	Q	I	H
S	Q	E	E	P	P	I	I	D	Q	L	G	L	G	V	S	P	G	Y	Q	T	F
V	S	C	Q	Q	Q	Q	L	S	F	L	P	P	P	W	G	D	C	S	S	A	S
L	N	P	N	Y	E	P	E	P	S	D	P	L	G	S	P	S	P	S	P	S	P
P	Y	T	L	M	G	C	R	L	A	C	E	T	R	Y	V	A	R	K	C	G	C
R	M	V	Y	M	P	G	D	V	P	V	C	S	P	Q	Q	Y	K	N	C	A	H
P	A	I	D	A	M	L	R	K	D	S	C	A	C	P	N	P	C	A	S	T	R
Y	A	K	E	L	S	M	V	R	I	P	S	R	A	A	A	R	F	L	A	R	K
L	N	R	S	E	A	Y	I	A	E	N	V	L	A	L	D	I	F	F	E	A	L
N	Y	E	T	V	E	Q	K	K	A	Y	E	M	S	E	L	L	G	D	I	G	G
Q	M	G	L	F	I	G	A	S	L	L	T	I	L	E	I	L	D	Y	L	C	E
V	F	R	D	K	V	L	G	Y	F	W	N	R	Q	H	S	Q	R	H	S	S	T
N	L	L	Q	E	G	L	G	S	H	R	T	Q	V	P	H	L	S	L	G	P	R
P	P	T	P	P	C	A	V	T	K	T	L	S	A	S	H	R	T	C	Y	L	V
T	Q	L																			